



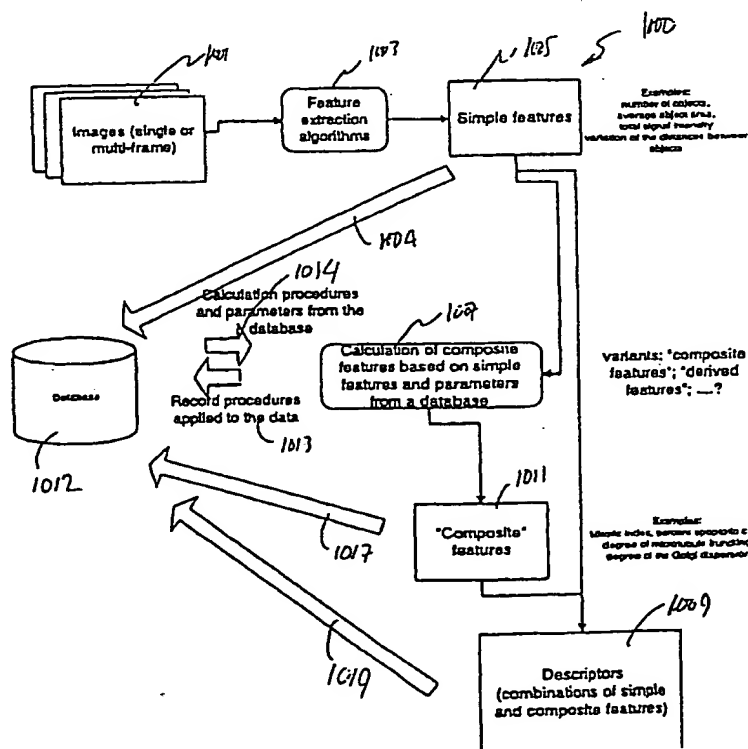
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : G06F 19/00		A2	(11) International Publication Number: WO 00/70528
			(43) International Publication Date: 23 November 2000 (23.11.00)
(21) International Application Number: PCT/US00/13154		(74) Agent: LOUIE, Michael, L.; Beyer Weaver & Thomas, LLP, P.O. Box 130, Mountain View, CA 94042-0130 (US).	
(22) International Filing Date: 12 May 2000 (12.05.00)		(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(30) Priority Data: 09/310,879 14 May 1999 (14.05.99) US 09/311,890 14 May 1999 (14.05.99) US 09/311,996 14 May 1999 (14.05.99) US 60/134,104 14 May 1999 (14.05.99) US Not furnished 26 April 2000 (26.04.00) US			
(71) Applicant: CYTOKINETICS, INC. [US/US]; Suite 2, 280 East Grand Avenue, South San Francisco, CA 94080 (US).			
(72) Inventors: SABRY, James, H.; 52 Buena Vista Terrace, San Francisco, CA 94117 (US). ADAMS, Cynthia, L.; 615 Georgia Avenue, Palo Alto, CA 94306 (US). VAISBERG, Eugeni, A.; 647 Pegasus Lane, Foster City, CA 94404 (US). CROMPTON, Anne, M.; 2 Bellair Place, San Francisco, CA 94133 (US). BLUM, Robert, I.; 17 Shoreview Avenue, San Francisco, CA 94121 (US). OESTREICHER, Donald, R.; 904 Old Town Court, Cupertino, CA 95014-4024 (US). SIGAL, Nolan, H.; 941 Berry Avenue, Los Altos, CA 94024 (US).		Published Without international search report and to be republished upon receipt of that report.	

(54) Title: METHOD AND APPARATUS FOR PREDICTIVE CELLULAR BIOINFORMATICS

(57) Abstract

Techniques for using information technology in therapeutics or drug discovery. In an exemplary embodiment, techniques for determining information about the properties of substances based upon information about structure of living or non-living cells exposed to substances are provided. A method according to the present invention enables researchers and/or scientists to identify promising candidates in the search for new and better medicines or treatments using, for example, a cellular informatics database. The present invention further teaches a system for acquiring knowledge from cellular information. The system has a database 1012 comprising a database management module ("DBMS"). The system also has a variety of modules, including a population module coupled to the DBMS for categorizing and storing a plurality of features (e.g., cell size, distance between cells, cell population, cell type) from an image acquisition device into the database. The system has a translation module coupled to the DBMS for defining a descriptor from a set of selected features from the plurality of features. In a specific embodiment, the descriptor is for a known or unknown compound, e.g., drug. A prediction module is coupled to the DBMS for selecting one of a plurality of descriptors from known and unknown compounds from the database based upon a selected descriptor from a selected compound. The selected compound may be one that is useful for treatment of human beings or the like.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**PATENT APPLICATION
METHOD AND APPARATUS FOR
PREDICTIVE CELLULAR BIOINFORMATICS**

COPYRIGHT NOTICE

5 A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever. The present description includes examples of
10 computer codes, which may be used to implement aspects of the present invention. Assignee of the present invention reserves all rights with respect to these codes and provides notice herein. Notice is hereby given © Cytokinetics, Inc. 1999.

BACKGROUND OF THE INVENTION

15 The present invention provides techniques for information management using a database platform. More particularly, the present invention provides a system including computer code that couples to a database device. The system provides for image capturing of living, dead, or fixed cells or cell fractions used to identify information about substances used on the cells or information about the cells themselves. Accordingly, the present invention can enable researchers and
20 scientists to identify promising candidates in the search for new and better medicines, for example, in drug discovery and development. The principles enumerated herein may, with equal facility, be applied to other applications, including but not limited to use in environmental applications such as determining chemical toxicities and other non-pharmaceutical toxicology uses.

25 For a long time, researchers in the pharmaceutical field have sought for better ways of searching for substances possessing properties that make them suitable as medicines. In the early days, researchers generally relied upon extracts from plants, dyes, and microbiological extracts for such substances. Examples of such substances include the pain reliever aspirin, the anti-cancer drug paclitaxel (brand
30 name TaxolTM), and the heart medication called digoxin. The number of useful medicines has generally been limited.

Purified substances having desirable bio-active properties are also often difficult to discover. Advances in traditional organic chemistry and more recently the rapid chemical synthesis methods often referred to as combinatorial chemistry have increased the number of compounds that researchers test for biological activity. Originally, substances were often initially tested on animals or humans to determine their biological activity. While results from such tests may identify a good drug candidate, they are often time consuming and costly, thus a limited number of substances can be tested. Therefore, pharmaceutical companies have turned to testing their ever-increasing libraries of substances against isolated proteins (drug targets) in biochemical assays that can be carried out at high throughput and low cost. It should be noted that the substances need to be tested in numerous protein tests, each customized for a particular drug target. Therefore, although each protein test may be run at a high-throughput, the design of multiple protein tests can be time-consuming. Substances deemed promising based on results from the protein tests are then tested in lower throughput cellular and animal tests.

There have been some attempts to use image acquisition techniques to screen a large number of substances based upon biological cell information. One such attempt is described in International Application No. WO 98/38490 in the names of Dunlay, et al. Dunlay et al. generally describes a conventional image acquisition system. This conventional system collects and saves images based on certain criteria that are predefined, not on a fixed area of an imaging surface. Additionally, the conventional system has poor lighting design, which makes image processing for multiple cells difficult. Furthermore, the conventional system is not designed for capturing, populating and utilizing a large database design. The conventional system is designed for customized cellular assays, not as a tool for generation of a cellular informatics database. Without such database capabilities the conventional system cannot be used for screening, analyzing, and comparing large quantities of cells from multiple experiments on multiple days in a predictive, efficient and cost effective manner.

What is needed is a rapid assay to assess the activity of compounds against multiple drug targets simultaneously in a cellular context. What is also needed are techniques for finding the effects of substances on cell function based upon searching and analyzing cellular information.

SUMMARY OF THE INVENTION

According to at least one embodiment of the present invention, techniques for determining information about effects of potential substances on cells are provided. In another exemplary embodiment, the present invention provides a novel system including hardware, computer codes, user interfaces, and a database for acquiring, storing and retrieving cellular and substance information. The cells can include living, dead, or fixed cells or fractions of cells. The present invention enables, *inter alia*, researchers and/or scientists to identify promising candidates in the search for new and better medicines or treatments using, for example, a cellular informatics database.

According to the present invention, a computer program for identification and verification of biological properties of substances can include code that causes a sample of a substance to be administered to a cell. The code determines one or more features for two or more cell components, or markers, in the presence of the substance. The code can form one or more descriptors from the features. Descriptors can be formed by combining features of two or more cell components as identified using the markers. The code can then search one or more descriptors obtained from prior administered substances upon cells in order to locate descriptors having a relationship to the descriptors noted for the substance under study. The code predicts properties of the administered substance based upon the properties of the prior administered substances using the relationship between the descriptors. The code can provide for identifying properties of substances based upon effects on cell characteristics. Candidate drug mechanisms of action, potency, specificity, pharmacodynamic, and pharmacokinetic parameters, toxicity, and the like can be used as substance properties.

In a specific embodiment, the present invention provides a system for acquiring knowledge from cellular information. The system has a database comprising a database management module ("DBMS"). The system also has a variety of other modules, including a population module that is coupled to the DBMS and serves to categorize and store a plurality of features (including but not limited to cell size, distance between cells, cell population, as well as sub-cellular features such as organelle location, protein location and sub-cellular constituent location and

movement) from an image acquisition device into the database. The system has a translation module coupled to the DBMS for defining a descriptor from a set of selected features from the plurality of features. In a specific embodiment, the descriptor is for a known or unknown compound, e.g., drug. A prediction module is coupled to the DBMS for selecting one of a plurality of a descriptors from known and unknown compounds from the database based upon a selected descriptor from a selected compound. The selected compound may be one that is useful for treatment of human beings or the like.

In a specific embodiment, the present invention provides a system for populating a database with cellular information. The system includes a cell holder (e.g., multi-well plate, chip, microfluidic assembly, or other cell chamber) comprising a plurality of sites in a spatial orientation. Each of the sites is capable of holding a plurality of cells to be imaged. Note – the light guide is one embodiment, but we don't want to be limited to it.

According to one embodiment, the present system also has an illumination apparatus including a liquid light guide operably coupled to the imaging device for highlighting the plurality of cells in a relatively even spatial manner for image capturing and measurement purposes. Still further, the liquid light guide allows sub-elements (e.g., filter, lamp) of the illumination apparatus to be placed at a remote location to prevent mechanical interference of the cell holder during image capturing. Alternative lighting methodologies may, with equal facility, be implemented.

The system also has an image-capturing device (e.g., charge coupled device camera, translation stage, shutter, microscope, software, shutter control) coupled to a computing device (e.g., computer, network computer, work station, analog computing device, on-board image-processor, and laptop). The image-capturing device is adapted to capture at least one image in at least one of the plurality of sites. One some embodiments, multiple images can be captured, where each image represents a different cell component (or portion). The image-capturing device can be adapted to convert the image into a digital representation, which highlights the feature or features of the one site.

A database storage device (e.g., relational database, object oriented database, mixed object oriented database) includes a database management element. The

database is coupled to the image capturing device. In a specific embodiment, the present system includes modules for feature extraction, generation of descriptions, and data preparation and analysis.

In a specific embodiment, the present invention provides a novel
5 system for determining an effect of a manipulation of a cell using one or more image frames. The system has a plate comprising a plurality of sites in a spatial orientation. Each of the sites is capable of holding a plurality of cells to be imaged. The system also has an image capturing device to capture a plurality of images of at least one site from the plurality of sites. The image capturing device is coupled to the computing
10 device. The system also has an image processing device to combine the plurality of images of at least one site or plurality of sites. The image processing device is operably coupled to the plate. An image processing device is also included. The image processing device can be adapted to form a digitized representation of the plurality of images from the site or plurality of sites. Furthermore, the system has a
15 database storage device comprising a database management element. The database can be adapted to retrieve the descriptor or descriptors of the plurality of features from the computing processing device and storing them in a selected manner.

In a specific embodiment, the present invention provides a system for capturing cellular information. The system also has an image acquisition system
20 comprising a charged coupled device camera adapted to capture an image of a plurality of manipulated cells in various stages of the cell cycle. The stages of the cell cycle are currently understood to include interphase, G0 phase, G1 phase, S phase, G2 phase, M phase, prophase, prometaphase, metaphase, anaphase, and telophase. The principles of the present invention specifically contemplate the application thereof on
25 additional cell cycle stages when and if they are identified.

An optical source is coupled to the image acquisition system for highlighting the plurality of manipulated cells in the various stages of the cell cycle. The illumination apparatus provides for an acquisition of the image of the plurality of manipulated cells. In a specific embodiment, the illumination apparatus has a liquid
30 light guide coupled to a light source at a remote location.

A variety of user interfaces are utile for accessing the several features of the present invention. Those having ordinary skill in the art will appreciate that different user interfaces may be required to support different research scenarios. The

present invention specifically contemplates the utilization of a wide variety of user interfaces.

Numerous benefits are achieved by way of the present invention over conventional techniques. The present invention can provide techniques for predictive cellular bioinformatics that can streamline a number of important decisions made in the drug discovery industry. The present invention can be implemented using off the shelf hardware including databases. In other aspects, the present invention can find useful information about substances as well as cells or portions of cells. Furthermore, the present invention can acquire more than one feature using more than one manipulation. Moreover, the present invention can provide information about a wide variety of cellular information that is not conventionally available. This information includes information about different cell components, e.g., nuclei and Golgi apparatus. Still further, the present invention provides an automated or semi-automated technique for acquiring images and populating a database. The present database can be combined with others such as genomics, and the like. Moreover, the present invention can be implemented to predict, *inter alia*, a mechanism of action, toxicity, target validation, and pre-clinical disease model.

A further understanding of the nature and advantages of the invention herein may be realized by reference to the remaining sections of the specification and the attached drawings.

BRIEF DESCRIPTION OF THE DRAWING

For more complete understanding of the present invention, reference is
5 made to the accompanying Drawing in the following Detailed Description of the
Invention. In the drawing:

Fig. 1 is a simplified system diagram according to an embodiment
according to the present invention;

10 Figs. 1A-1B are more detailed diagrams of database systems according
to embodiments of the present invention;

Fig. 2 is a simplified block diagram according to an alternative
embodiment according to the present invention;

Figs. 3-6 are simplified diagrams of system elements according to
embodiments of the present invention;

15 Figs. 7A-7K illustrate representative block diagrams of simplified
process steps in a particular embodiment according to the present invention;

Fig. 8A-8F illustrate representative quantified descriptors of effects of
manipulations on images of cells in a particular experiment;

20 Fig. 9 illustrates example images for different types of morphologies in
a particular experiment;

Fig. 10 illustrates a distribution of various morphologies in a cell
population responsive to drug concentration in a particular experiment;

Fig. 11 illustrates a graph of quantified features of effects of
manipulations on cells in a particular experiment;

25 Fig. 12 illustrates effects of external agents on cells in a particular
experiment;

Fig. 13 illustrates 4 panels for each marker for a plurality of A549 cells
in a particular experiment;

30 Fig. 14 illustrates 4 panels for each marker for a plurality of OVCAR-3
cells in a particular experiment;

Fig. 15 illustrates 4 panels for each marker for a plurality of OVCAR-3
cells at 20x in a particular experiment;

Fig. 16 illustrates 4 panels for each marker for a plurality of OVCAR-3 cells at 40x in a particular experiment;

Fig. 17 illustrates a representative input for a morphometric analysis program in a particular embodiment according to the present invention; and

5 Figs. 18-19 illustrate examples of the generation of pseudo-sequences and clustering in a particular embodiment according to the present invention.

Fig. 20 is a block diagram for a first research scenario;

Fig. 21 is a block diagram for a second research scenario; and

Fig. 22 is a block diagram for a third research scenario.

10 Reference numbers refer to the same or equivalent parts of the invention throughout the several figures of the Drawing.

DETAILED DESCRIPTION OF THE INVENTION

According to the present invention, techniques for determining information about manipulated cells or substances based upon living, fixed, or dead cell structures or portions of cells are provided. In an exemplary embodiment, the present invention provides a novel system including computer codes coupled to a database and user interfaces for acquiring, storing and retrieving such information. Other embodiments provide a novel image capturing system for providing digitized representations of live and dead cell structures or the like.

Fig. 1 is a simplified system diagram 10 of a cellular knowledge-based system according to an embodiment of the present invention. This diagram is merely an example and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. The present system 10 includes a variety of elements such as a computing device 13, which is coupled to an image processor 15 and is coupled to a database 21. The image processor receives information from an image capturing device 17, which image processor and image capturing device are collectively referred to as the imaging system herein. The image capturing device obtains information from a plate 19, which includes a plurality of sites for cells. These cells can be biological cells that are living, fixed, dead, cell fractions, cells in a tissue, and the like. The computing device retrieves the information, which has been digitized, from the image processing device and stores such information into the database. A user interface device 11, which can be a personal computer, a work station, a network computer, a personal digital assistant, or the like, is coupled to the computing device.

Fig. 1A is a simplified diagram of a database system 1000 according to an embodiment of the present invention. This diagram is merely an example and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. Database system 1000 includes a variety of techniques for processing images from biological cells, e.g., fixed, living, and dead cells, and cell portions. As shown, images are acquired 1001. These images can be from a single frame or multiple frames. As merely an example, an image processing system may analyze such images. One example of

such an image processing system is described below, but should not be construed as limiting certain claims.

In a specific embodiment, cell samples are manipulated using a compound (e.g., substance, drug). The cell samples are imaged for a simple portion or portions, e.g., manipulated cell substructure, manipulated spatial feature of cell, cell density. Image processing techniques are used to extract the feature or features from the image or images. The features can be an independent or a dependent set of cell characteristics (which may be predominately visual) including, for example, count, area, perimeter, length, breadth, fiber length, fiber breadth, shape factor, elliptical form factor, inner radius, outer radius, mean radius, equivalent radius, equivalent sphere volume, equivalent prolate volume, equivalent oblate volume, equivalent sphere surface, average intensity, total intensity, optical density, radial dispersion, texture difference, and others. Each of these features corresponds to a similar manipulation by a compound. Each manipulation forms a new set of features, which are identifiable to the compound. Once each set of features has been extracted, the feature set is populated into a database. Accordingly, the database includes many sets of features, where each set corresponds to a different manipulation for a selected cell. Each set of features corresponding to a manipulation provides a descriptor, which is also stored in the database. The descriptor is a "finger print" including each feature for the manipulation. Each descriptor may be unique, or may have similarities to other descriptors or may even be the same as other descriptors for known and unknown manipulations.

The present system retrieves features, which we define as simple features herein, and forms composite features from them. More than one feature can be combined in a variety of different ways to form these composite features. In particular, the composite feature can be any function or combination of a simple feature and other composite features. The function can be algebraic, logical, sinusoidal, logarithmic, linear, hyperbolic, statistical, and the like. Alternatively, more than one simple feature can be combined in a functional manner (e.g., arithmetic, algebraic). As merely an example, the composite feature equals a sum of feature 1 and feature 2, where these features correspond to the same manipulation. Alternatively, the composite feature equals feature 1 divided by feature 2. Alternatively, the composite feature equals feature 1 minus feature 2. Alternatively,

the composite feature equals a constant times feature 1 plus feature 2. Of course, there are many ways that the composite feature can be defined. The present system also stores 1017 these features in the database. The composite features can also be further combined with simple features. Once these features are defined as descriptors, they are stored 1019 in the database.

Fig. 1B is a simplified diagram of a database system engine 2000 according to an embodiment of the present invention. This diagram is merely an example and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize many other variations, modifications, and alternatives. The engine can be implemented into the present database for populating, searching, and predicting compound or cell characteristics. As merely an example, engine 2001 includes an input/output module 2008. The input/output module is used to input and output information from the database. The information includes, among others, a plurality of feature sets, which correspond to many manipulations. Additionally, the information includes descriptors, which each corresponds to a set of features from the manipulation. The database also has a population module, which is used to configure the features based upon an entity relationship, which has been predetermined.

The database engine also has other modules. In particular, the database has a transcription module, which transfers a preselected set of features and creates a descriptor from them. The transcription module can be used to take a known compound, which has features, to transcribe them into a descriptor. Alternatively, the transcription module can be used to take an unknown compound, which has features, to transcribe them into a descriptor. These descriptors are provided into the database for subsequent use. Finally, the database engine has a prediction module, which can be used to potentially predict a property (e.g., mechanism of action) of an unknown compound. Here, the unknown compound is provided with a descriptor, but the property of the compound is unknown. In one embodiment, the prediction module compares a descriptor of an unknown compound with the many descriptors of known compounds, which were in the populated database. Depending upon the matching criteria, the prediction module will attempt to uncover one or more descriptors of known compounds. Once the prediction module finds the descriptors of the known compounds based upon the descriptor for the unknown compound, it identifies a potential property of such unknown compound for analysis and review. Here, it is

believed that certain features of the known compound, which are similar to those features of the unknown compound may uncover a property to the unknown compound. Details of the present software engine are described more fully below.

Fig. 2 is a simplified block diagram 20 of a cellular knowledge-based system according to an alternative embodiment of the present invention. This diagram is merely an example and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. Like reference numerals are used in the present diagram as the previous diagram for easy cross-referencing, but are not intended to be limiting in any manner.

The present diagram 20 includes a variety of elements such as a processor 13 or computing device coupled to a database 11. The processor can be used for retrieving and storing information from the database. The system also includes a plurality of system elements, such as a cleaner 23, a dispenser 25, and an image capturing system 27, which are also coupled to the database in some embodiments. These elements can be coupled to each other through a network or the like. As merely an example, the network can be a NetWareTM network from Novell Corporation or an internet network or the Internet but can also be others and any combination thereof. The system also has an output device 31, which can be used to output information from the database, processor, or other system elements. Details of these elements are described more fully below in reference to the Figs.

Figs. 3-5 are simplified drawings of system elements according to embodiments of the present invention. These diagrams are merely examples and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. As merely an example,

Fig. 3 is a simplified diagram of a processor or computing device 13. The computing device 13 includes a bus 112 which interconnects major subsystems such as a central processor 114, a system memory 116 (e.g., random access memory), an input/output ("I/O") controller 118, an external device such as a display screen 124 via a display adapter 126, a keyboard 132 and a mouse 146 via an I/O controller 118, a SCSI host adapter (not shown), and a floppy disk drive 136 operative to receive a floppy disk 138.

The computing device has other features. Storage Interface 134 may act as a storage interface to a fixed disk drive 144 or a CD-ROM player 140 operative

to receive a CD-ROM 142. Fixed disk 144 may be a part of computing device or may be separate and accessed through other interface systems. A network interface 148 may provide a direct connection to a remote server via a telephone link or to the Internet. Network interface 148 may also connect to a local area network ("LAN") or other network interconnecting many computer systems. Many other devices or subsystems (not shown) may be connected in a similar manner. Also, it is not necessary for all of the devices shown in Fig. 3 to be present to practice the present invention, as discussed below. The devices and subsystems may be interconnected in different ways from that shown in Fig. 3. The operation of a computer system such as that shown in Fig. 3 is readily known in the art and is not discussed in detail in this application. Computer code to implement the present invention, may be operably disposed or stored in computer-readable storage media such as system memory 116, fixed disk 144, CD-ROM 140, or floppy disk 138. The computer code can be organized in terms of processes or modules, depending upon the application. That is, the computer code can include a prediction module, a translation module, or other modules to carry out the functionality described herein, as well as others.

Figs. 4 and 5 are simplified diagrams of an imaging system 200 according to an embodiment of the present invention. As shown, the imaging system 200 includes a variety of features such as housing 203, which holds a stage assembly 204. The stage assembly includes an x-stage movement element 206, which is along an x-direction, and a y-stage movement element 207, which is along a y-direction. The imaging system also includes a z-direction movement element, which is perpendicular to the x-y plane. The z-direction movement motor can be attached to the stage, or to the objective nosepiece by way of the microscope housing, or as an external motor between the objective and the microscope housing. The stage can align in any one of the directions to an accuracy of one micron and less, or one-half micron and less, or one-quarter micron and less, depending upon the embodiment.

The stage holds a plate 202 or cell holder, which houses one of a plurality of samples. The plate includes a spatial array 209 of process sites. Each of the process sites can include a plurality of cells and solutions depending upon the embodiment. Each of the sites can carry a sufficient amount of solution to prevent substantial evaporation of the sample during processing in some embodiments. In embodiments for large scale analysis, the plate includes at least 96 sites, or more than

or equal to 384 sites, or more than or equal to 1,536 sites. The plate bottom is transparent and thin, which allows light to pass through the sample. Additionally, the plate is made of a suitable chemical resistant material. As merely an example, the plate can be either a 96, or 384, or 1536 or other formats from places such as Becton Dickinson of Franklin Lakes, NJ, or Corning Science Products of Corning, NY. In a preferred embodiment, the plate is a Corning Costar black-walled 96 well plate catalog #3904 from Corning Science Products of Corning, NY, but should not be limited to these in some applications, but can be others.

Also shown is the condenser for the microscope 201, which can be used to collect phase, DIC, or bright field images of the cells. Images resulting from the illumination of the samples to fluorescence, phase, DIC, or bright field techniques are collected using an image capturing device 208, which captures an image or images of cells from the plate. In a specific embodiment, the microscope is an inverted configuration with the objectives on the bottom of the plate and the condenser disposed overlying an upper surface of the sites, while the image capturing device underlies the sites. Images captured by the imaging device, whether analogue or digital, are viewed by a monitor or other devices. The image capturing device can be any camera assembly such as a charge coupled device camera, which is known as a CCD camera, or other high resolution camera capable of capturing images from the sites. In a specific embodiment, the camera is an interline CCD camera which does not require an external shutter.

In a specific embodiment, the present imaging system can be any suitable unit that is flexible for automated image collection using multi-well plastic plates. The imaging system also should be adapted to collect high-resolution images of cells on plastic or glass plates, cell growth chambers, or coverslips. The system also can be used for imaging multiple cell markers in multiple imaging conditions. To accomplish this, the microscope system has a variety of elements such as a light source, a motorized excitation filter wheel and shutter, x-y-z-motorized stage, excitation and emission filters, Fluor phase and DIC objectives, motorized objective nosepiece, dichroic filters, motorized dichroic filter cubes, phase and DIC rings and prisms, CCD camera, and software control. As merely an example, the present imaging system can have components such as those listed in the Table below.

DESCRIPTION	MAKER	MODEL
Microscope	Zeiss	100M
(x-y) motorized stage	Prior	
Xenon lamp	Sutter	Lambda
Filter wheel	Sutter	Lambda-10
Microtitre Plate holder	Prior	500-H223R
Isolation Table	Kinetic Systems	9101-24-85
Objective Spacers	Polytec PI	P-721.90
Camera	Hamamatsu	C47-95
Computer	IBM	IntelliStation
Software	Metamorph	v.4
Objectives	Zeiss	Achroplan 10x/0.25 LD-Achroplan 20x/0.4 LD-Achroplan 40x/0.6

Table: Image Acquisition System Elements

5 In a specific embodiment, the present system has the following capabilities, which are not intended to be limiting.

Image acquisition

1) Ability to automatically acquire multi-wavelength images from multiple sites on one multi-well plate, to sequentially name image files, and to log any
10 imaging parameter information with image files.

2) Ability to link images with a larger database/spreadsheet of information.

3) Ability to automatically collect multiple plates by interfacing the imaging system with a robotic arm.

15

X-Y control

1) Ability to place 96, 384, or 1536 well plates onto microscope stage and move to each well sequentially.

2) Ability to return to each well and collect another round of images (multi-site time-lapse) or ability to collect rapid time-lapse information at each well (time-lapse of many wells).

3) Ability to collect a low magnification image, automatically determine features which may be of interest, automatically change the objective to a higher magnification, and collect high magnification images of a fixed number of those identified cells in the sample.

4) Ability to collect multiple frames in each site.

10 Z control

1. Ability to auto-focus with substantially minimal damage to biological specimen or fluorophore.

2. Ability to auto-focus rapidly.

15 The present embodiment of the imaging system is shown by way of Figs. 5A and 5B. These diagrams are merely examples and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. The present imaging system 40 includes a variety of elements such as a microscope 41, which is preferably an epi-fluorescent microscope, but can be confocal, multiphoton, or hybrid types. The microscope includes elements 20 41A, the motorized Z-axis; 41B, the motorized dichroic filter cube holder; and 41C, the motorized objective nosepiece. In one embodiment, the microscope is a Model 100M made by Zeiss. The microscope communicates to computer 51 through control lines 73, 75, and 76. The imaging system also has camera 50 coupled to controller 25 50A and computing device 51, which oversees and controls operations of the elements of the imaging system.

The present microscope includes drivers for spatially moving a stage in two dimensions, including an x-direction, a y-direction, and moving the objective nosepiece in a z-direction in a Cartesian coordinate system. The z-direction 30 movement is provided using a fast z-motor, which can make z-direction adjustments within a predetermined time. The z-direction movement generally provides for focussing of the sample to the camera. The focussing occurs within the predetermined time of preferably ten seconds and less, or five seconds and less, or one

second and less, depending upon the embodiment. As merely an example, the z-motor or positioner can be a model PIFOC objective nanopositioner made by a company called Physik Instrumente of Waldbronn, Germany, but also can be others. The z-motor couples to computer 51 through line 63, which may also include a
5 controller. Depending upon the embodiment, a second z-motor 41A connected to the computer 51 by line 73 may be used to keep the z-motor 42 in the center of its travel. Alternatively, in other embodiments the stage could be provided with a z-motor allowing for movement of the stage in the z-direction.

The present stage also includes an x-y stage 43. The x-y stage moves
10 plate 59, e.g., 96 site, 384 site, 1536 site. The x-y stage moves plate in an x-y spatial manner. The stage has an accuracy or repeatability of about 1 micron and less, or about 2 microns and less. The stage can move in a continuous manner or a stepped manner. The stage also can move up to 30 mm/sec. or faster. The stage also can move 1 mm/sec. and less, depending upon the embodiment. The stage can also step
15 0.1 micron and less or 1 micron and less, as well as other spatial dimensions. The stage can be one such as a Proscan Series made by Prior Scientific of Rockland, MA but can also be others. The stage is controlled via control line 61 through controller 43A, which couples to computer 51 through control line 65.

The stage includes plate holder 44. The plate holder can hold a single
20 plate. In other embodiments, plate holder can also hold multiple plates. The plate holder can use mechanical, electrical, fluid, vacuum and other means for holding the plate or plates. The plate holder also is sufficiently stable for securing the plate. As merely an example, the plate holder is a Model 500-H223R made by Prior Scientific of Rockland, MA. In some embodiments, the plate holder may need adjustment in
25 the z-direction to provide for a desirable focus of a sample on a plate. In these embodiments, the plate holder is supported by spacers 45 or a plurality of stage pins, which mechanically elevate the plate holder in the z-direction. These pins are generally made of a suitable material for supporting such plate holder and also are sufficiently resistant to chemicals and the like.

30 In some embodiments, the entire imaging system is placed on an isolation table 49. The isolation table is disposed between the microscope and support structure. The isolation table is designed to prevent excessive vibration of the plate. The isolation table is made of a suitable material such as steel and honeycomb but can

be others. The table has a thickness of about 8 inches or preferably less than about 24 inches. In one embodiment, the table is Model 9101-24-85 made by Kinetic Systems of Boston, MA.

The imaging system also has a lamp or illumination assembly 62. The lamp assembly provides for a light source (See reference letter B) to a plurality of elements in the imaging system. For easy reading, the light path is defined by the dotted lines, which are not intended to be limiting. The lamp assembly has a variety of elements such as a Xenon lamp 46. The Xenon lamp provides light at about 320 to 700 nanometers (Prefocused). The Xenon lamp is 175 or 300 Watts. As merely an example, the lamp can be a Lambda Model made by Sutter Instrument Company of Novato, CA.

Referring to Fig. 5B, the lamp assembly also has a cold mirror 58, an excitation filter wheel 48, excitation filter(s) 55, and an excitation light shutter 57. As shown, light is derived from the Xenon lamp, reflects off of the cold mirror 58; traverses through the excitation filter or filters 55, and is controlled by the excitation light shutter 57. The lamp assembly has filter wheel 48, which houses one of a plurality of filters, including excitation filters. The shutter and filter wheel are controlled via control lines 67, which are coupled to a computer 51 or other type of computing device. The control lines 67 are coupled through controller 57A (for element 57) and controller 48A (for element 48) via control line 69 to computer 51.

Preferably, light traverses from the lamp assembly through a light guide 47 to illuminate features within the plate. The light guide is suitably selected to have a flexible member, which can be used to place lamp source at a remote location away from the imaging device. The flexible member substantially keeps any vibration from the lamp assembly away from the imaging device. In some embodiments, the member is at least 1 foot away from the imaging device. The light guide is a guide, which is a flexible hose-type sleeve. The sleeve is filled with a liquid such as an aqueous solution containing chloride or phosphate. A thin layer may be formed on the inside of the sleeve. The layer can be a containing tetrafluoroethylene and hexafluoropropylene, or containing tetrafluoroethylene and perfluoromethyl vinyl ether, or tetrafluoroethylene and perfluoropropyl vinyl ether. An example of such a light guide is described in International Application No. WO/98/38537 filed February 29, 1997, and assigned to NATH, Gunther. The liquid

light guide has less than about 30% transmission loss of the light at a remote location such as the imaging system.

Light is derived from the lamp assembly and directs off of filter 56, which directs the light upward. Filter 56 can be a dichroic and emission filter, as well as others. The light traverses through microscope nosepiece 41C, and traverses through objective spacers 54. An objective 53 magnifies the light toward a predetermined point on the plate 59. The objective can be, for example, made by Zeiss of Jena, Germany, as well as other companies. The objective can be one of a plurality including 1X, 10X, 20X, 40X, and others, depending upon the application. Magnification can be further expanded or contracted by intermediate optics between the objective and the camera. Selection of filter or filters is controlled by computer 51 via control line 75.

The camera 50 captures an image of cells from plate 59. The image is obtained from light scattering off of cells or portions of cells in the plate through objective 53, through objective spacers, through filters 56, which are captured at camera 50. In this preferred embodiment, the camera is a digital camera, but can be an analogue camera. The digital camera is a CCD camera, which has 1280 by 1024 pixels, or more or less. The pixels can be 6.7 microns in dimension or more or less. The camera preferably is substantially free from an external shutter to quickly capture a plurality of images of cells from the plate. The camera is controlled via control line 71 through controller 50A, which connects to computer 51 through control line 70. The present invention can also include other types of image acquisition devices selected from at least an epifluorescence, a confocal, a total-internal reflection, a phase, a Hoffman, a bright field, a dark field, a differential interference contrast, an interference reflection, or multi-photon illumination device.

The present imaging system stores images on a high density memory device 60. The high density memory device is preferably optical, but can also be magnetic. The high density memory device can be any suitable unit that is capable of storing a plurality of images from a plurality of sites in the plate. The memory device can be a compact disk, which would generally use a compact disk burner or the like. Depending upon the embodiment, the high density memory device is used to archive the images that are captured from the camera in the imaging system. Further details

of the imaging system can be found throughout the present specification, and more particularly below.

As merely an example, the present invention can be implemented using the following sequence of steps, which have been described in a journal entry form.

- 5 Here, images are opened and objects are identified based on a background value that has been edited in starting image acquisition. Information is maintained in a spreadsheet or other database format, which has the following information for each object:

Image Name	Image Plane	Image Date and Time
Elapsed Time	Object #	Total area
Pixel area	Area	Hole area
Relative hole area	Standard area count	Perimeter
Length	Breadth	Fiber length
Fiber breadth	Shape factor	Ell. form factor
Inner radius	Outer radius	Mean radius
Average gray value	Total gray value	Optical density
Radial dispersion	Texture Difference Moment	EFA Harmonic 2, Semi-Major Axis
EFA Harmonic 2, Semi-Minor Axis	EFA Harmonic 2, Semi-Major Axis Angle	EFA Harmonic 2, Ellipse Area
EFA Harmonic 2, Axial Ratio	EFA Harmonic 3, Semi-Minor Axis	

10

After computations are done, the log file is saved. In particular, the file is saved in an appropriate place with an appropriate name.

In a specific embodiment, the present invention provides the following detailed example of journal entries, which should not limit the scope of the invention.

Set Up Sequential File Names	Interactive: user sets up prefix name and image storage directory
Open Data Log	Opens a DDE (Excel) File
Annotate Log File	Interactive: experimental information that will go into the first line of the log file of stage positions
Stage (Go to Origin)	Origin is set as the center of well A1
Stage (Move to Absolute Position)	Offset to upper left hand corner of well (1410, 1621)
Stage (Log Position)	
Stage (Scan Wells)	User picks wells to scan: runs 3x3 image collection.jnl.

3X3 IMAGE COLLECTION.jnl

Stage (Scan)	Takes 9 images of well, -1600 motor steps apart from left to right 3 columns and 3 rows, runs FOCUS, COLLECT IMAGE, SAVE SEQUENTIAL FILE NAME.JNL.
--------------	--

5

FOCUS, COLLECT IMAGE, SAVE SEQUENTIAL FILE NAME.jnl.

Stage (Log Position)	Logs stage position of each image
ADC – Focus	Opens up the manual focusing window with whatever focus time is current set
Show Message and Wait	Interactive: user hits enter to continue when done focusing

ADC-Acquire from Digital Camera	Takes Hoechst image
Save Using Sequential File Names	
Close	Closes image window

START IMAGE ANALYSIS.jnl

Low Pass	3x3 convolution of already opened image
Low Pass	3x3
Show Region Statistics	Interactive: Show entire image statistics. Calculate background subtraction value for step 4. by: INTENSITY Average + INTENSITY Std. Dev.
Arithmetic	Interactive: User inputs subtraction value from 3. into the constant Value field
Threshold image	Creates threshold 1 unit above 0 to 4096
Integrated Morphometry – Load State	Loads Start Image Analysis.ima Classifier 100 < area < 200000
Integrated Morphometry – Measure	Interactive: Shows area summary information about all objects. The average number is used as the Standard Area in 8.
Object Standards - Set Object Standards	Interactive: User inputs average area value from 7. into Standard Area box to be used by automated IMA for all images

IMA OBJECTS.jnl

Low Pass	3x3 convolution
Low Pass	3x3 convolution
Arithmetic	This background subtraction value needs to be manually entered into this journal from the value determined in START IMAGE ANALYSIS.jnl step 3
Threshold Image	1 unit above 0
Integrated Morphometry – Load State	Hoecsht.IMA Classifier 200 < area < 200000
Integrated Morphometry – Measure	Measures statistical info for all objects
Run Journal	Runs log obj and sum data.jnl

Log obj and sum data.jnl

Integrated Morphometry – Log Data	Logs object data into Sheet 1
Integrated Morphometry – Log Data	Log summary data into Sheet 2

5

COLLECT AUTOMATED IMA DATA IN ONE SPREADSHEET.jnl

Run Journal	Runs OPEN OBJECT LOG DDE FILE.JNL
Loop for all Images in a Directory	Loops IMA OBJECTS.jnl
Close Summary Log	
Close Object Log	User must manually save Excel spreadsheet

OPEN OBJECT LOG DDE FILE.jnl

Open Object Log	Opens a DDE object log into sheet 1 of an Excel spreadsheet
Open Summary Log	Opens a summary log into sheet 2

COLLECT AUTOMATED IMA DATA IN ONE SPREADSHEET 16 BIT
IMAGES.jnl

Arithmetic	Interactive: Opens Arithmetic window for user to input background subtraction level from START IMAGE ANALYSIS.jnl step 3
Run Journal	Runs OPEN OBJECT LOG DDE FILE.JNL
Loop for all Images in a Directory	Interactive: Runs IMA OBJECTS 16 bit.jnl. User picks directory from which to choose.

5

IMA OBJECTS 16bit.jnl

Low Pass	3x3 convolution
Low Pass	3x3 convolution
Copy to 8-bit Image	No autoscale, to new untitled image
Save Using Sequential File Name	Saves 8bit image using previously defined Sequential File names.
Arithmetic	This background subtraction value needs to be manually entered into this journal from the value determined in START IMAGE ANALYSIS 16 TO 8 BIT.jnl step 5
Threshold Image	1 unit above 0 to 255

Integrated Morphometry – Load State	Hoecsht.IMA Classifier 200 < area < 200000
Integrated Morphometry – Measure	Measures statistical info for all objects
Run Journal	Runs log obj and sum data.jnl

START IMAGE ANALYSIS 16 to 8 BIT.jnl

Copy to 8-bit Image	No autoscale, to new untitled image
Close	Closes 16 bit image
Low Pass	3x3 convolution
Low Pass	3x3 convolution
Show Region Statistics	Interactive: Show entire image statistics. Calculate background subtraction value for step 6. by: INTENSITY Average + INTENSITY Std. Dev.
Arithmetic	Interactive: User inputs subtraction value from 5. into the constant Value field
Threshold image	Creates threshold by 1 unit above 0 to 255
Integrated Morphometry – Load State	Loads Start Image Analysis.ima Classifier 100 < area < 200000
Integrated Morphometry – Measure	Interactive: Shows area summary information about all objects. The average number is used as the Standard Area in 10.
Object Standards - Set Object Standards	Interactive: User inputs average area value from 9. into Standard Area box to be used by automated IMA for all images

IMA OBJECTS WITH NEW LOG FILE.jnl

Run Journal	OPEN OBJECT LOG DDE FILE.JNL
Run Journal	IMA OBJECTS.jnl
Close Summary Log	
Close Object Log	User must manually save every Excel spreadsheet generated.

INTERACTIVE IMA OBJECTS.jnl

Threshold Image	User manually sets threshold
Integrated Morphometry – Load State	Hoechst.IMA Classifier 200 < area < 200000
Integrated Morphometry – Measure	Objects
Integrated Morphometry – Log Data	Into open object.log file

5

COLLECT INTERACTIVE IMA DATA.jnl

Close Object Lo g	
Open Object Log	Interactive
Annotate Log File	Interactive: experimental information that will go into the first line of the object log file
Loop for all Images in Directory	Runs INTERACTIVE IMA OBJECTS.jnl

CHANGE FILTER, COLLECT IMAGE, SAVE SEQUENTIAL FILE
NAME.jnl

Stage (Log Position)	
ADC-Focus	

Show Message and Wait	Interactive – user presses Enter when done focusing
ADC – Acquire from Digital Camera	Hoechst
Save Using Sequential File Name	
Close	Close open image

COLLECT HOECHST AND FITC.jnl

Run Journal	FOCUS, COLLECT IMAGE, SAVE SEQUENTIAL FILE NAME.JNL
Run Journal	CHANGE FILTER, COLLECT IMAGE, SAVE SEQUENTIAL FILE NAME.jnl

3X3 IMAGE COLLECTION HOECHST FITC.jnl

Stage (Scan)	COLLECT HOECHST AND FITC.jnl
--------------	------------------------------

5

AUTOMATED 3X3 IMAGE COLLECTION HOECHST FITC.jnl

Set Up Sequential File Names	Interactive: user sets up prefix name and image storage directory
Open Data Log	Excel DDL files
Annotate Log File	Interactive: experimental information that will go into the first line of the log file of stage positions
Stage (Go to Origin)	Origin is set as the center of well A1
Stage (Move to Absolute Position)	Offset to upper left hand corner of well (1410, 1621)

Stage (Log Position)	
Stage (Scan Wells)	Interactive: user picks wells to scan: runs 3X3 IMAGE COLLECTION HOECHST FITC.jnl

AUTOMATED IMAGE COLLECTION.jnl

Set Up Sequential File Names	Interactive: user sets up prefix name and image storage directory
Open Data Log	Opens a DDE (Excel) File
Annotate Log File	Interactive: experimental information that will go into the first line of the log file of stage positions
Stage (Go to Origin)	Origin is set as the center of well A1
Stage (Log Position)	
Stage (Scan Wells)	Interactive: user picks wells to scan: runs FOCUS, COLLECT IMAGE, SAVE SEQUENTIAL FILE NAME.JNL. Well to well travel = (-9035, -9035)

5

STARTUP.jnl

Install and Configure Devices	Open Stage Meta Devices
Set Live Video Channel	

Preferences	<u>Measure Objects</u> : Draw failed classifier objects, Exclude objects that touch the edge of the image, Enable Elliptical Fourier Parameters, turn off Warn users when measurement data will be erased <u>Image Saving</u> : Save Tiff/stk using LZW compression <u>Image Windows</u> : Use transparent thresholds.
Configure Default Paths	C:\Metamorph Data C:\Metamorph Data\Common Settings
Load Journal Taskbar	Common.JTB

Nested Journals

Automated 3x3 Image Collection

5 *Loop* 3x3 image collection
 Loop focus, collect image, save sequential file name

Automated 3x3 image collection Hoechst FITC

10 *Loop* 3x3 image collection Hoechst FITC
 loop Collect Hoechst and FITC
 focus, collect image, save sequential file name
 change filter, collect image, save sequential file name

Automated image collection

15 *Loop* focus, collect image, save sequential file name

Collect automated IMA data in one Spreadsheet

Open object log DDE file

Loop IMA objects

Log obj and sum data

Collect automated IMA data in one spreadsheet 16 bit images

5 Open object log DDE file

Loop IMA objects 16 bit

Log obj and sum data

Although the above has been generally described in terms of a specific
10 user interface and software code, other user interfaces and code can also be used. One
of ordinary skill in the art would recognize many other variations, alternatives, and
modifications.

Fig. 6 is a simplified diagram 600 of a cleaning and dispensing system
according to an embodiment of the present invention. This system 600 includes a
15 variety of elements such as a dispensing head 609, which is coupled to a plurality of
pipettes 601. The pipettes input and output fluids or solutions from plate 603. The
plate has a plurality of sites, each of which can be used to input cells or a combination
of cells and solution. The system also has elements to house solutions 605, which are
used to manipulate cell samples in the plate. The dispensing head is supported
20 through a support member 607, which is sufficiently rigid to allow for movement of
the head. The dispenser is coupled to the present system in a mechanical and
electrical manner, which provides for a fully integrated system for providing cell
samples to the imaging system according to the present invention.

Fig. 7A illustrates a representative block flow diagram of simplified
25 process steps of a method for determining properties of a manipulation based upon
effects of the manipulation on one or more portions of one or more cells in a
particular embodiment according to the present invention. This diagram is merely an
illustration and should not limit the scope of the claims herein. One of ordinary skill
in the art would recognize other variations, modifications, and alternatives. In step
30 700, one or more samples of cells can be provided. These cells can be live, dead, or
fixed cells, or cell fractions. The cells also can be in one of many cell cycle stages,
including G0, G1, S, G2 or M phase, M phase including the following cell cycle
stages: interphase, prophase, prometaphase, metaphase, anaphase, and telophase.

Cell components tracked in presently preferable embodiments can include proteins, protein modifications, genetically manipulated proteins, exogenous proteins, enzymatic activities, nucleic acids, lipids, carbohydrates, organic and inorganic ion concentrations, sub-cellular structures, organelles, plasma membrane, adhesion complex, ion channels, ion pumps, integral membrane proteins, cell surface receptors, G-protein coupled receptors, tyrosine kinase receptors, nuclear membrane receptors, ECM binding complexes, endocytotic machinery, exocytotic machinery, lysosomes, peroxisomes, vacuoles, mitochondria, Golgi apparatus, cytoskeletal filament network, endoplasmic reticulum, nuclear membrane, proteosome apparatus, chromatin, nucleolus, cytoplasm, cytoplasmic signaling apparatus, microbe specializations and plant specializations.

The following table illustrates some markers and cell components commonly used by embodiments according to the present invention. Other markers can be used in various embodiments without departing from the scope of the invention.

Cell component	Marker	Disease State
Plasma membrane (including overall cell shape)	Carbocyanine dyes Phosphatidylserine Various lipids Glycoproteins	Apoptosis-Cancer Apoptosis-Neural degenerative Ds
Adhesion complexes	Cadherins Integrins Occludin Gap junction ERM proteins CAMs Catenins Desmosomes	Thrombosis Metastasis Wound healing Inflammatory Ds Dermatologic Ds
Ion Channels and Pumps	Na/K Atpase Calcium channels Serotonin reuptake pump CFTR	Cystic fibrosis Depression Congestive Heart Failure Epilepsy

G coupled receptors	β adrenergic receptor Angiotensin receptor	Hypertension Heart Failure Angina
Tyrosine kinase receptors	PDGF receptor FGF receptor IGF receptor	Cancer Wound healing Angiogenesis Cerebrovascular Ds
ECM binding complexes	Dystroglycan Syndecan	Muscular Dystrophy
Endocytotic machinery	Clathrin Adaptor proteins COPs Presenilins Dynamin	Alzheimer's Ds
Exocytotic machinery	SNAREs Vesicles	Epilepsy Tetanus Systemic Inflammation Allergic Reactions
Lysosomes	Acid phosphatase Transferrin	Viral diseases
Peroxisomes/Vacuoles		Neural degenerative Ds
Mitochondria	Caspases Apoptosis inducing factor F1 ATPase Fluorescein Cyclo-oxygenase	Apoptosis Neural degenerative Ds Mitochondrial Cytopathies Inflammatory Ds
Golgi Apparatus	Lens Culinaris DiOC6 carbocyanine dye COPs	

Cytoskeletal Filament Networks	Microtubules Actin Intermediate Filaments Kinesin, dynein, myosin Microtubule associated proteins Actin binding proteins Rac/Rho Keratins	Cancer Neural degenerative Ds Inflammatory Ds Cardiovascular Ds Skin Ds
Endoplasmic Reticulum	SNARE PDI Ribosomes	Neural degenerative Ds
Nuclear Membrane	Lamins Nuclear Pore Complex	Cancer
Proteosome Apparatus	Ubiquityl transferases	Cancer
Chromatin	DNA Histone proteins Histone deacetylases Telomerases	Cancer Aging
Nucleolus	Phase markers	
Cytoplasm	Intermediary Metabolic Enzymes BRCA1	Cancer
Cytoplasmic Signaling Apparatus	Calcium Camp PKC pH	Cardiovascular Ds Migraine Apoptosis Cancer
Microbe Specializations	Flagella Cilia Cell Wall components: Chitin synthase	Infectious Ds

Plant specializations	Choloroplast Cell Wall components	Crop Protection
-----------------------	--------------------------------------	-----------------

Then, in a step 702, one or more samples of the manipulation can be provided to the cells. Manipulations can comprise one or any combination of chemical, biological, mechanical, thermal, electromagnetic, gravitational, nuclear, or temporal factors, for example. For example, manipulations could include exposure to chemical compounds, including compounds of known biological activity such as therapeutics or drugs, or also compounds of unknown biological activity. Or exposure to biologics that may or may not be used as drugs such as hormones, growth factors, antibodies, or extracellular matrix components. Or exposure to biologics such as infective materials such as viruses that may be naturally occurring viruses or viruses engineered to express exogenous genes at various levels. Bioengineered viruses are one example of manipulations via gene transfer. Other means of gene transfer are well known in the art and include but are not limited to electroporation, calcium phosphate precipitation, and lipid-based transfection. Manipulations could also include delivery of antisense polynucleotides by similar means as gene transfection. Other genetic manipulations include gene knock-outs or gene mutations. Manipulations also could include cell fusion. Physical manipulations could include exposing cells to shear stress under different rates of fluid flow, exposure of cells to different temperatures, exposure of cells to vacuum or positive pressure, or exposure of cells to sonication. Manipulations could also include applying centrifugal force. Manipulations could also include changes in gravitational force, including sub-gravitation (the preferred embodiment in outer space). Manipulations could include application of a constant or pulsed electrical current. Manipulations could also include irradiation. Manipulations could also include photobleaching which in some embodiments may include prior addition of a substance that would specifically mark areas to be photobleached by subsequent light exposure. In addition, these types of manipulations may be varied as to time of exposure, or cells could be subjected to multiple manipulations in various combinations and orders of addition. Of course, the type of manipulation used depends upon the application.

Then, in a step 704, one or more descriptors of a state in the portions of the cells in the presence of the manipulation can be determined using the images

collected on the imaging system. Descriptors can comprise scalar or vector values, representing quantities such as area, perimeter, dimensions, intensity, gray level, aspect ratios, and the like. Other types of descriptors include, but are not limited to, one or any combination of characteristics such as a cell count, an area, a perimeter, a length, a breadth, a fiber length, a fiber breadth, a shape factor, a elliptical form factor, an inner radius, an outer radius, a mean radius, an equivalent radius, an equivalent sphere volume, an equivalent prolate volume, an equivalent oblate volume, an equivalent sphere surface area, an average intensity, a total intensity, and an optical density. These descriptors can be average or standard deviation values, or frequency statistics from the descriptors collected across a population of cells. These descriptors can be further reduced using other methods such as principal component analysis and the like. In some embodiments, the descriptors include features from different cell portions or cell types. That is, a first feature can be from a nuclei and a second feature is from another cell structure such as Golgi apparatus, mitochondria, spacing between cell structures or cells themselves, as well as many others.

A presently preferable embodiment uses descriptors selected from the following table. Other descriptors can also be used without departing from the scope of the invention.

Name of Parameter	Explanation/Comments
Count	Number of objects
Area	
Perimeter	
Length	X axis
Width	Y axis
Shape Factor	Measure of roundness of an object
Height	Z axis
Radius	
Distribution of Brightness	
Radius of Dispersion	Measure of how dispersed the marker is from its centroid
Centroid location	x-y position of center of mass
Number of holes in closed objects	Derivatives of this measurement might include, for

	example, Euler number (= number of objects - number of holes)
Elliptical Fourier Analysis (EFA)	Multiple frequencies that describe the shape of a closed object
Wavelet Analysis	As in EFA, but using wavelet transform
Interobject Orientation	Polar Coordinate analysis of relative location
Distribution Interobject Distances	Including statistical characteristics
Spectral Output	Measures the wavelength spectrum of the reporter dye. Includes FRET
Optical density	Absorbance of light
Phase density	Phase shifting of light
Reflection interference	Measure of the distance of the cell membrane from the surface of the substrate
1,2 and 3 dimensional Fourier Analysis	Spatial frequency analysis of non closed objects
1,2 and 3 dimensional Wavelet Analysis	Spatial frequency analysis of non closed objects
Eccentricity	The eccentricity of the ellipse that has the same second moments as the region. A measure of object elongation.
Long axis/Short Axis Length	Another measure of object elongation.
Convex perimeter	Perimeter of the smallest convex polygon surrounding an object
Convex area	Area of the smallest convex polygon surrounding an object
Solidity	Ratio of polygon bounding box area to object area.
Extent	proportion of pixels in the bounding box that are also in the region
Granularity	
Pattern matching	Significance of similarity to reference pattern
Volume measurements	As above, but adding a z axis

Then, in a step 705, a database of cell information can be provided.

Next, in a step 706, a plurality of descriptors can be searched from a database of cell information in order to locate descriptors based upon one of the descriptors of the manipulation. Then, in a step 708, properties of the manipulation are predicted based upon the properties of the located descriptors. Properties can comprise toxicity, specificity against a subset of tumors, mechanisms of chemical activity, mechanisms of biological activity, structure, adverse biological effects, biological pathways, clinical effects, cellular availability, pharmacological availability, pharmacodynamic properties, clinical uses and indications, pharmacological properties, such as absorption, excretion, distribution, metabolism and the like.

In a particular embodiment, step 706 comprises determining matching descriptors in the database corresponding to a prior administration of the manipulation to the descriptors of the present administration of the manipulation. In a particular embodiment according to the present invention, combinations of measurements of scalar values can provide predictive information. A database can be provided having one or more "cellular fingerprints" comprised of descriptors of cell-substance interactions of drugs having known mechanisms of action with cells. Such descriptors can be analyzed, classified, and compared using a plurality of techniques, such as statistical classification and clustering, heuristic classification techniques, a technique of creating "phylogenetic trees" based on various distance measures between descriptors from various drugs. In this embodiment, numeric values for the descriptors can be used by comparison techniques. A phylogenetic tree can be created that illustrates a statistical significance of the similarity between descriptors for the drugs in the database. Because the drugs used to build the initial database are of known mechanism, it can be determined whether a particular scalar value in a descriptor is statistically predictive. Finally, a compound descriptor with no known mechanism of action can be queried against the database and be statistically compared and classified among the drugs in the database that the compound most resembles.

In a particular embodiment, relationships between measured morphological properties of images and physiological conditions can be determined. Relationships can include, for example, treatment of different cell lines with chemical compounds, or comparing cells from a patient with control cells, and the like. In a presently preferable embodiment, comparisons can be performed on acquired image

features. Some embodiments can comprise statistical and neural network - based approaches to perform comparisons of various features. The foregoing is provided as merely an example, and is not intended to limit the scope of the present invention. Other techniques can be included for different types of data.

5 In some embodiments, classification, clustering and other types of predictive data analysis can be performed on features extracted from cell images. In a presently preferable embodiment, statistical procedures for comparisons, classification and clustering are performed on data obtained from imaging cells.

Fragments of data preparation and pre-formatting (S language):

```
10       >tmp.frame <- Generic.Summary  
      >names1 <- paste("Cell.line.5", tmp.names, sep=".")  
      > by.compound.matrix <- as.matrix(arranged.by.compound)
```

Example of the code for principal component analysis (data
15 preparation) using S language:

```
      all.data.princomp <- menuPrincomp(data =  
      by.compound.matrix, scores = T, cor = "Correlation",  
      na.action = T, print.short = T, print.importance = T,  
      print.loadings = T, cutoff.loadings = 0.1,  
20   plot.screplot = T, plot.loadings = T, plot.biplot = T,  
      plot.biplot.choices = c(1,2), predict.p = F)
```

Example of clustering using a divisive hierarchical clustering
algorithm:

```
25       > div.hier.2.manhattan.cluster$call  
      diana(x = tmp.sum.by.comp, diss = F, metric =  
      "manhattan",  
          stand = T, save.x = T, save.diss = T)
```

30 Another embodiment utilizes existing tools for biological sequence similarity searches, classification, and phylogenetic analysis. In a particular embodiment, numbers in a numerical descriptor can be substituted by one or more of nucleic acid or amino acid codes according to a one of several sets of rules. Once

converted into a corresponding nucleotide or amino acid sequence representation, the fingerprints can be analyzed and compared using software and algorithms known in the art for genetic and peptide sequence comparisons, such as GCG, a product of Genetics Computer Group, with company headquarters in Madison WI. Select
5 embodiments comprising such approaches enable the use of a broad array of sophisticated algorithms to compare, analyze, and cluster gene and protein sequences. Many programs performing this task are known to those of ordinary skill in the art, such as for example, the PHYLIP (PHYlogeny Interference Package) a package of programs for inferring phylogenies (evolutionary trees) described in (Feldenstein, J.
10 1996 Methods Enzymol 266:418-427 and Feldenstein, J. 1981 J. Mol. Evol. 17(6):368-376).

Embodiments can perform such analysis based upon factors such as numerical value, statistical properties, relationships with other values, and the like. Further details of a step of manipulation are noted more particular below.

15 Fig. 7B illustrates a representative block flow diagram of simplified process steps for determining one or more descriptors of a state in the portions of the cells in the presence of the manipulation of step 704 of Fig. 7A in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill
20 in the art would recognize other variations, modifications, and alternatives. In a step 712, an image of a cell portion is obtained. In some embodiments, the cell portion is visualized with a fluorescently labeled marker that is specific for the portion or portions of interest. A cell portion can include, for example, one or more of the following: nuclei, Golgi apparatus, and other features. The cell portion may vary in
25 select embodiments according to the invention. Then, in a step 714, a digitized representation of the image obtained in step 712 is determined. In some embodiments, steps 714 and step 712 can comprise a single step. These embodiments use a digital imaging means such as a digital camera, to obtain a digital image of the target directly. Next, in a step 716, the digital representation of the image is
30 processed to obtain image features. Image features can include such quantities as area, perimeter, dimensions, intensity, aspect ratios, and the like. Then, in a step 718 descriptors can be determined from the image features. Descriptors can comprise scalar or vector quantities and can comprise the image features themselves, as well as

composed features, such as shape factor derived by a relationship $4\pi * \text{area} / \text{perimeter}$, and the like. Descriptors can also comprise statistical quantities relating to feature characteristics across a population of cells, such as a standard deviation, and average, and the like.

5 In a preferred embodiment, cells can be placed onto a microscope, such as a Zeiss microscope, or its equivalent as known in the art. A starting point, named Site A01, is identified to the microscope. A plurality of exposure parameters can be optimized for automated image collection and analysis. The microscope can automatically move to a new well, automatically focus, collect one or more images, at
10 one or more wavelengths, move to a next well, and repeat this process for all designated wells in a multiple well plate and for multiple plates. A file having a size and an intensity distribution measurement for each color and rank for each well can then be created for the images acquired. Based on this information, a user or a computer can revisit sites of interest to collect more data, if desired, or to verify
15 automated analysis. In a presently preferred embodiment, image automatic focus and acquisition can be done using computer software controlling the internal Z-motor of the microscope. Images are taken using a 10x, 20x, or 40x air long working distance objectives. Sometimes multiple images are collected per well. Image exposure times can be optimized for each fluorescent marker and cell line. The same exposure time
20 can be used for each cell line and fluorescent marker to acquire data.

Fig. 7C illustrates a representative block flow diagram of simplified process steps for obtaining images of cell portions of step 712 of Fig. 7B in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill
25 in the art would recognize other variations, modifications, and alternatives. The method is generally outlined by the steps below:

- (1). In a step 720, a sample is provided to the imaging device. Samples can be provided in 96 well plates and the like. The sample may be loaded into a microscope, such as a Zeiss microscope or equivalent.
- 30 (2). In a step 722, a set of optical filters is selected to shine light of the appropriate wavelength to illuminate the first sample, which may be contained in a first well designated A01.

(3). In a step 724, an automatic focusing procedure is performed for the site. In a particular embodiment, the internal z-motor of the microscope which is attached to the objective nosepiece is used for automatic focusing of the microscope. In an alternative embodiments, the plate holding the samples is moved to perform automatic focusing of the microscope, or focusing can be performed by moving optical components attached to the microscope and the like.

(4). In a step 726, images are collected for the site. Images can be collected for every color at every site. Present embodiments can provide images for up to four colors. However, embodiments are contemplated that can provide more colors by using either a monochromator coupled with excitation filters which are on a filter wheel, or by digitally separating overlapping fluorophores. Those knowledgeable in the field will know that given calibration images of single fluorophores, a look-up table can be devised which will allow for the digital removal of fluorescence bleed-through of fluorescence which may occur in optical channels other than the one for which that filter has been optimized in instances of using more than one fluorophore at once. Cell growth and density information is also collected. Cell density is determined by what percentage of the area being imaged is inhabited by cells. In some embodiments, imaging can be facilitated using one or more biosensors, molecules such as non-proteins, i.e., lipids and the like, that are luminescently tagged. However, some embodiments can also use fluorescence polarization and the like. Fluorescence polarization is a homogeneous fluorescence technology where the excited state of the molecule lasts much longer than in normal fluorescence, taking seconds to minutes to reach equilibrium, obliterating the need to wash away fluorescence markers that are not specifically bound to a marker. Further, embodiments can detect differences in spectral shifts of luminescent markers. Some fluorescence markers, such as Nile Red sold by Molecular Probes of Eugene, OR, will change its emission peak wavelength depending on its environment. One can detect these changes by monitoring the level of fluorescence at both wavelengths and reading out at ratio of the two.

(5). In a step 728, a determination is made whether more fields of view need to be taken for a particular color. If this is so, then processing continues at step 726 at a new site. Otherwise, processing continues with a decisional step 730.

Images can now be taken by repeating step 726. In a preferred embodiment 4 to 9 images are collected at each site.

(5). In a step 730, a determination is made whether more optical configurations need to be taken in order to obtain images for all differently-marked cell portions the sample. If this is so, then in a step 732 a new optical configuration is determined. Images for the new optical configuration can now be taken by repeating steps 726 and 728.

(6). In a decisional step 734, after all optical configurations and images for fields of view in a sample have been obtained, a determination is made whether any further samples remain to be analyzed. If so, a new sample is brought into view and processing continues with step 720. Otherwise, image processing is complete. In a presently preferable embodiment, image data can be stored on a CD ROM using a CD ROM burner, such as CRW4416 made by Yamaha of Japan. However, other mass storage media can also be used.

Fig. 7D illustrates a representative block flow diagram of simplified process steps for processing digitized representations of step 716 of Fig. 7B in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. The method is generally outlined by the steps below:

(1). In a step 740, a digitized image input is preprocessed . Preprocessing might include, but is not limited to, such operations as background subtraction, thresholding, smoothing, adoptive filtering, edge enhancements, contrast enhancements, histogram equalization. A particular combination of preprocessing steps can be applied to images in successive steps or in parallel to copies of the image.

A simplified example of a smoothing and background subtraction procedure in a MatLab language is presented in computer code below:

```
function Isubtracted = cmBackgrSubtrl(I,k)
30 % cmBackgrSubtrl(I,k) - simple flat background (=modal*k)
subtraction
% Y = cmBackgrSubtrl(I, k) - image Y is generated by
```



```

    % subtraction (with saturation) of modal pixel value of I
    multiplied by k
    % DEFAULT - k=1
    %
5   if (nargin == 1)
        k=1;
    end
    if (size(k)~=1)
        error('cmBackgrSubtr1: parameter k should be a number.
10  Exiting...');
    end

    %modpixnum = floor(size(I(:),1)/2);
    %sortedval = sort( double(I(:)) );
15  %modpixel = sortedval(modpixnum);
    modpixel = median(double(I(:)));
    bg = k*modpixel;

    Isubtracted = mmsubm( uint8(I), uint8(round(ones(
20  size(I))*k*modpixel )) );

```

An example of a procedure for thresholding in computer code (MatLab) is presented below:

```

function thresh = GetThreshByPerim1(I, M)
25 % GetThreshByPerim1(I) Finds optimal thresholding value
    for image I
    % N = GetThreshByPerim1(I) Finds thresholding value N for
    image I
    % N = GetThreshByPerim1(I, M) - tests threshold values up
30  to M
    % DEFAULT M = maximum pixel value in I
    % note that GetThreshByArea is significantly faster
    % finds a threshold value that causes the maximal change
    in the

```

```
% total perimeter of the objects (Russ ????)
% see Matlab_Auto_threshold1_1-23-99.doc for more details
% Note: works somewhat better on SMOOTH images (i.e.
medfilt2(I, [3 3]) two times

5
if (nargin == 0)
    error (strcat( mfilename, ' : at least one parameter
required'));
elseif (nargin == 1)
10     M = double(max(I(:)));      %test thresholds up to
maximum pixel value in I
elseif (nargin > 2)
    error (strcat (mfilename, ' : too many parameters'));
end

15
if (size(M)>1)
    error (strcat(mfilename, ' : argument M should be a
number'));
end

20
Minval = double( min(I(:)));
step = 1;

%generate vertical vector perims with total perimeters of
25 objects at different
%threshold values
for i=Minval : step : M
    bwI = im2bw(I, i/255);
    prI = bwperim(bwI);
30     pr = sum(prI(:));
    if (exist('perims', 'var') == 0) %perims is yet
undefined
        perims = pr;
    else
```

```

        perims = cat(1, perims, pr);
    end
end

5  % vector prdiffs contains differences between successive
    perimeters
    prdiffs = diff(perims);
    mindecrease = min(prdiffs);
    minvalues = find(prdiffs == mindecrease);
10  index_of_mindecrease = minvalues(1);
    thresh = index_of_mindecrease + 1;

    % =====end GetThresh1=====

15      Thresholding provides a specific intensity, such that pixels darker than
the threshold are deemed black, and pixels lighter than the threshold are considered
white. The thresholded image can be processed using binary image processing
techniques in order to extract regions.

        (2). In a step 742+744, the digitized image input is subjected to object
20  identification. This can be accomplished by a variety of procedures, for example by
thresholding or edge detection and subsequent morphological opening and closing.
Edge detection can be accomplished by means of gradient-based or zero-crossing
methods, such as Sobel, Canny, Laplassian, Perwitt, and other methods.

        An example of object identification procedure based on Canny edge
25  detection (in MatLab language) is presented below:

function Imask = cmMaskDNA1(I);
% cmMaskDNA1 - generates binary mask for cell nuclei
through edge detection
30 % Imask = cmMaskDNA1(I)
% PARAMETERS
%     I - intensity image (grayscale)
% OUTPUT
%     Imask - BW image with objects from I

```

```

%
% For more details see Notebook Matlab_DNA_masking1_1-22-
99.doc
% Uses SDC Morphology Toolbox V0.7

5
if (nargin ~= 1)
    error('Wrong number of input parameters');
end
if (nargout ~= 1)
10    error('Wrong number of output parameters: one output
    argument should be provided');
end

15 Imask = edge(I, 'canny');
Imask = mmdil(Imask, mmsecross(1));
Imask = mmero ( mmclohole(Imask,mmsecross(1)));
Imask = mmedgeoff(Imask, mmsecross(1));
% note that mmedgeoff this command removed FILLED OBJECTS
20 but not touching OUTLINES.
% these outlines can be removed by filtering:
Imask = medfilt2(Imask, [5 5]);

%=====end cmMaskDNA1
25 =====

```

However, embodiments can also use other techniques, such as Fast Fourier Transforms (FFT) and the like as known in the art without departing from the scope of the present invention.

30 (3). In a step 746, a plurality of region features can be determined. For example, in a representative embodiment, image features can include such quantities as area, perimeter, dimensions, intensity, aspect ratios, and the like. Features not directly related to individual objects are also being extracted.

An example of a procedure for extraction of some of the features (MatLab language) is presented below:

```

function OData = cmGetObjectsData(I, Ilabel)
5  % cmGetObjectsData returns array measurements of objects
  in image "I" masked by "Ilabel"
  % EV 2-3-99; 2-10-99
  % OData = cmGetObjectsData(I, Ilabel) returns an array of
  morphological and intensity measurements
10 %   taken from a grayscale image "I". Objects are
  identified on a mask image Ilabel, usually
  %   created by bwlabel()
  % OUTPUT:
  % Each row in the output array OData represents
15 individual object
  % columns contain the following measurements:
  %
  %   1 - Index ("number" of an object);      8 -
  Solidity;
20 %   2 - X coordinate of the center of mass; 9 - Extent;
  %   3 - Y coordinate      "-"; 10 - Total
  Intensity;
  %   4 - Total Area (in pixels);              11 - Avg.
  Intensity;
25 %   5 - Ratio of MajorAxis/MinorAxis;        12 - Median
  Intensity;
  %   6 - Eccentricity;                        13 - Intensity of
  20% bright pixel
  %   7 - EquivDiameter;                      14 - Intensity of
30 80% bright pixel
  %
  % For details on morphological parameters see information
  on MatLab imfeature();

```

```
% Intensity parameters are either obvious or are
documented in comments in this file.

if (nargin ~= 2)
5   error ('function requires exactly 2 parameters');
end
if (nargout ~= 1)
    error ('function has 1 output argument (array X by
14)');
10 end

% finished checking arguments

% first collect morphological parameters in a structure
15 array:
ImStats = imfeature(Ilabel, 'Area', 'Centroid',
'MajorAxisLength',...
    'MinorAxisLength', 'Eccentricity', 'EquivDiameter',
...
20 'Solidity', 'Extent', 8 );

% now convert it into array (matrix) while collecting
intensity data for each object:

25 %preallocate output array:
numobjects = size(ImStats, 1);
OData = zeros(numobjects, 14);
%now convert ImStats into array and add intensity data to
it
30 for k=1:numobjects
    OData(k, 1) = k;
    OData(k, 2) = ImStats(k).Centroid(1);
    OData(k, 3) = ImStats(k).Centroid(2);
    OData(k, 4) = ImStats(k).Area;
```

```
OData(k, 5) = (ImStats(k).MajorAxisLength) /  
(ImStats(k).MinorAxisLength);  
OData(k, 6) = ImStats(k).Eccentricity ;  
OData(k, 7) = ImStats(k).EquivDiameter;  
5 OData(k, 8) = ImStats(k).Solidity;  
OData(k, 9) = ImStats(k).Extent;  
  
% now collect and assign intensity parameters from  
image I  
10  
    object_pixels = find( Ilabel == k);  
    object_area = size(object_pixels, 1); %same as total  
number of pixels in the object  
    object_intensities = double(I(object_pixels)); %  
15 need to convert to double to do math  
    sorted_intensities = sort(object_intensities); %  
will need to get median, 20% and 80% pixels  
    total_intensity = sum(object_intensities, 1);  
    avg_intensity = total_intensity / object_area;  
20 median_intensity = sorted_intensities( floor(  
object_area/2 ) + 1 );  
    pix20 = sorted_intensities( floor(object_area*0.2)+1  
) ; %brightest pixel among dimmest 20%  
    pix80 = sorted_intensities( floor(object_area*0.8)+1  
25 ) ;  
  
OData(k, 10) = total_intensity;  
OData(k, 11) = avg_intensity;  
OData(k, 12) = median_intensity;  
30 OData(k, 13) = pix20; %brightest pixel among dimmest  
20%  
OData(k, 14) = pix80; %dimmest pixel among brightest  
20%  
end %for
```

```
%===== end function
cmGetObjectsData() =====
```

- 5 (4). In a step 748, quantitative descriptors, characterizing cell state are calculated based on the feature measurements extracted at step 746. For example, histogram distribution of intensities of cell nuclei provides information about the population cell cycle stages.

In a particular embodiment according to the present invention, data analysis techniques for describing the fluorescence patterns of cell portions in multiple cell lines in the presence and absence of compounds are provided. Automated image analysis techniques can include determining one or more regions from around nuclei, individual cells, organelles, and the like, called "objects" using a thresholding function. Objects that reside on the edge of an image can be included or excluded in various embodiments. An average population information about an object can be determined and recorded into a database, which can comprise a database text file or Excel spreadsheet, for example. However, embodiments can use any recording means without departing from the scope of the present invention. Values measured can be compared to the visual image. One or more types of numerical descriptors can be generated from the values. For example, descriptors such as a number of objects, an average, a standard deviation of objects, a histogram (number or percentage of objects per bin, average, standard deviation), and the like can be determined.

In a particular embodiment according to the present invention, data can be analyzed using morphometric values derived from any of a plurality of techniques commonly known in the art. For example, a software package called MetaMorph Imaging System, provided by Universal Imaging Corporation, a company with headquarters in West Chester, PA and NIH Image, provided by Scion Corporation, a company with headquarters in Frederick, Maryland.

30 Fluorescent images can be described by numerical values, such as for example, an area, a fluorescence intensity, a population count, a radial dispersion, a perimeter, a length, and the like. Further, other values can be derived from such measurements. For example, a shape factor can be derived according to a relationship

$4\pi \times \text{area} / \text{perimeter}$. Other values can be used in various embodiments according to the present invention. Such values can be analyzed as average values and frequency distributions from a population of individual cells.

In a particular embodiment according to the present invention, techniques for the automatic identification of mitotic cells are provided. Image analysis techniques employing techniques such as multidimensional representations, frequency-based representations, multidimensional cluster analysis techniques and the like can be included in various embodiments without departing from the scope of the present invention. Techniques for performing such analyses are known in the art and include those embodied in MatLab software, produced by MathWorks, a company with headquarters in Natick, MA.

Scalar values providing efficacious descriptors of cell images can be identified using the techniques of the present invention to perform predictive analysis of drug behavior. In a presently preferred embodiment, a plurality of heterogeneous scalar values can be combined to provide descriptors for each manipulation. By applying predictive analysis routines to the collections of these descriptors, predictive information about any number of manipulations and cell interactions can be extracted.

Fig. 7E illustrates a representative block flow diagram of simplified process steps for analyzing image feature values to obtain descriptors of cell state of step 718 of Fig. 7B in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. Fig. 7E illustrates an input data of descriptors of known manipulations 319. A step 320 of reformatting and transforming data 319 to formats suitable for analysis is performed. Additionally, a "cleaning" process can eliminate outlying data points and the like in the data. Then, in a step 322, a decision is made whether to continue with step 324 or with step 326 based upon determining a particular type of analysis appropriate for the present application or particular type of prediction. If decisional step 322 determines processing should continue with step 324, then, in that step, an error estimate using a set of test descriptors is performed to estimate the quality of a prediction and processing continues with step 320. Once an optimal prediction is achieved, processing continues with step 326. In step 326, optimal transformation parameters and prediction methods are selected for use in

steps 328 and 330 which analyze data about an unknown manipulation. In a step 328, a solution is generated based upon any of techniques including training a neural network, solving a mathematical equation, applying decision tree rules and/or the like. In a step 330, an input data set of unknown descriptors 318 is reformatted and transformed based upon the optimal transformation parameters selected in step 326 using the transformation procedures in steps 320, 322 and 324. In a step 332, predictions techniques are applied to the reformatted manipulations from step 330 and the solution generated in step 328 and a plurality of properties of known manipulations 317 (e.g., therapeutic properties, and the like) in order to determine a prediction of properties of unknown manipulation 316.

Fig. 7F illustrates a representative block flow diagram of simplified process steps for a method of mapping a manipulation of cells to a physiological characteristic in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. The method is generally outlined by the steps below:

(1) In a step 750, a plurality of cells, e.g., dead, live, cell fractions or mixtures of cells are provided.

(2) Then, in a step 752, the plurality of cells is manipulated, where manipulation occurs using a source(s) from one or a combination selected from an electromagnetic, electrical, chemical, thermal, gravitational, nuclear, temporal, or a biological source.

(3) Next, in a step 754, a feature value is captured from the plurality of cells. The feature value can include one or any combination of characteristics such as cell count, area, perimeter, length, breadth, fiber length, fiber breadth, shape factor, elliptical form factor, inner radius, outer radius, mean radius, equivalent radius, equivalent sphere volume, equivalent prolate volume, equivalent oblate volume, equivalent sphere surface area, average intensity, total intensity, and optical density. This list is not meant to be limiting.

(4) Then, in a step 756, a degree of presence of one or more feature values is assigned for each manipulation.

(5) In a step 758, the feature values from the plurality of cells are stored in memory locations. From the memory locations the values can be used for

statistical analyses to produce predictive information about the relatedness of the descriptors of the manipulations to one another. This information is used to infer properties of the manipulations.

Fig. 7G illustrates a representative block flow diagram of a simplified process steps for a method for populating a database with manipulated biological cell information in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. The method is generally outlined by the steps below:

(1) In a step 760, a plurality of cells in various stages of the cell cycle, A montage image that was used as a source to generate data in Appendix A is presented in Fig. 12., such as for example, the stages of interphase, prophase, metaphase, anaphase, and telophase are provided.

(2) Then, in a step 762, each of the cells in the various stages of mitotic development is manipulated.

(3) Next, in a step 764, an image of the plurality of manipulated cells is captured using image acquisition techniques in order to provide a morphometric characteristic of each of the manipulated cells.

(4) As a preferable option, in a step 766, an image database may be populated with the image of the plurality of manipulated cells.

(5) Following step 764 or optional step 766, a morphological value is calculated from the image in a step 768.

(6) In a step 770, the database is populated with the morphological value.

Fig. 7H illustrates a representative block flow diagram of simplified process steps for a method for populating a database with manipulated biological information, e.g., image acquisition parameters, image feature summary information, and well experimental parameters in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. Fig. 7H illustrates a step 780 in which cells are placed into site on a plate and a manipulation is applied. Then, in a step 781 an image is taken of the cells. In step 782, the image is transferred to an image archive

database. Then, in a step 783, well experimental parameters are entered into the database 787. Well experimental parameters can include cell type, manipulation and the like. In a step 784, image acquisition parameters are transferred to database 787. Image acquisition parameters can include file name, fluorophores and the like. In a
5 step 785, the image acquired in step 781 is analyzed. Then, in step 786, an image feature summary from the analysis step 785 is transferred to database 787.

In step 788, a lookup table for all analyses is provided to database 787. The lookup table provides information about the analyses. In a step 789, a query of database 787 for process data is performed. The results are reformatted. Then in a
10 step 790, the database 787 is queried. Next, in a step 791, features of the manipulations stored in the database are combined and reduced. Next, in a step 793, reduced features of step 791 can be compared. In a step 792, the results of step 793 are recorded in database 787. Then, in a step 794, a report of predictions based on comparisons performed in step 793 is generated.

15 Fig. 7I illustrates a representative block flow diagram of simplified process steps for acquiring images of manipulated biological information, e.g., cells, cell tissues, and cell substituents in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations,
20 modifications, and alternatives. Fig. 7I illustrates a step 770 in which a user sets up an image analysis procedure. Then, in a step 772, an image is read into image analysis software. Next, in a step 774, patterns and objects are identified in the image using one or more algorithms. Next, in a step 776, sets of features are extracted from the image. Then, in a step 778, feature information, descriptor values and the like are
25 exported to the database, such as database 787 of Fig. 7H, for recording. Next, in a decisional step 779, a determination is made whether any more images should be taken. If this is so, processing continues with step 772. Otherwise, image acquisition processing is completed.

Fig. 7J illustrates a representative block flow diagram of simplified
30 process steps for populating, acquiring and analyzing images of manipulated biological information in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations,

modifications, and alternatives. Fig. 7J illustrates a step 300 of placing a plate onto an imaging stage and reading a bar code. Then, in a step 301 an autofocus procedure is performed. Next, in a step 302, a first optical filter configuration is selected and an image is collected. Then, in a decisional step 303, a determination is made whether more than one image per optical configuration can be taken. If so, then, in a step 304, a new position within the well is targeted and another image is collected. Then, in a decisional step 305, a determination is made whether any more images need to be collected. If this is so, step 304 is repeated until all images for a particular well have been collected. After one or more images are collected for the well, in a step 306, the stage is returned to a starting position within the well, and a montage is created from collected images. The results are named with a unique file name and stored.

In a decisional step 307, a determination is made whether any more optical channels in the well can be imaged. If this is so, then in a step 308 the next optical filter configuration is selected and an image is collected. Processing then continues with decisional step 303, as described above. Otherwise, if no further optical channels in the well can be imaged, then in a decisional step 309 a determination is made whether any wells remain to be imaged. If not all wells have been imaged, then in a step 310, the stage moves to the next well and processing continues with step 301, as described above. Otherwise, if all wells on the plate have been imaged, then in a decisional step 311, a determination is made whether any more plates can be processed. If this is so, then processing continues with step 300 as described above. Otherwise, in a step 312, the information is stored to a CD or other storage device as a backup.

Fig. 7K illustrates a representative block flow diagram of simplified process steps compound based upon information about effects of one or more known compounds on a cell population in a particular embodiment according to the present invention. This diagram is merely an illustration and should not limit the scope of the claims herein. One of ordinary skill in the art would recognize other variations, modifications, and alternatives. Fig. 7K illustrates a step 340 of populating a database with descriptors for known compounds. Such descriptors can be determined from imaging the cell population. However, in some embodiments, descriptors can be derived by measurements and combinations of measurements and the like. Then, in a step 342, descriptors for the unknown compound are determined from imaging a

second cell population. The second cell population has been treated with the unknown compound. Then, in a step 344, a relationship between the descriptors determined from the unknown compound with the descriptors determined from the known compounds can be determined. Finally, in a step 346, an inference can be made about the unknown compound based upon the descriptors of the known compounds from the relationship determined in step 344.

Accordingly, the present invention provides a novel database design. In a particular embodiment according to the present invention, a method for providing a database comprises measurement of a potentially large number of features of one or more sub-cellular morphometric markers. Markers can be from any of a large variety of normal and transformed cell lines from sources such as for example, human beings, fungi, or other species. The markers can be chosen to cover many areas of cell biology, such as, for example markers comprising the cytoskeleton of a cell. The cytoskeleton is one of a plurality of components that determine a cell's architecture, or "cytoarchitecture". A cytoarchitecture comprises structures that can mediate most cellular processes, such as cell growth and division, for example. Because the cytoskeleton is a dynamic structure, it provides a constant indication of the processes occurring within the cell. The cytoarchitecture of a cell can be quantified to produce a one or more scalar values corresponding to many possible cellular markers, such as cytoskeleton, organelles, signaling molecules, adhesion molecules and the like. Such quantification can be performed in the presence and absence of drugs, peptides, proteins, anti-sense oligonucleotides, antibodies, genetic alterations and the like. Scalar values obtained from such quantification can provide information about the shape and metabolic state of the cell.

In a presently preferred embodiment, scalar values can comprise morphometric, frequency, multi-dimensional parameters and the like, extracted from one or more fluorescence images taken from a number of cellular markers from a population of cells. Two or more such scalar values extracted from a plurality of cell lines and markers grown in the same condition together comprise a unique "fingerprint" or descriptor that can be incorporated into a database. Such cellular descriptors will change in the presence of drugs, peptides, proteins, antisense oligonucleotides, antibodies or genetic alterations. Such changes can be sufficiently unique to permit a correlation to be drawn between similar descriptors. Such

correlations can predict similar properties or characteristics with regard to mechanism of action, toxicity, animal model effectiveness, clinical trial effectiveness, patient responses and the like. In a presently preferred embodiment, a database can be built from a plurality of such descriptors from different cell lines, cellular markers, and compounds having known mechanisms of action (or structure, or gene response, or toxicity).

The present invention also provides database and descriptor comparisons according to other embodiments. In a particular embodiment according to the present invention, measurement of scalar values or features can provide predictive information. A database can be provided having one or more "cellular fingerprints" comprised of descriptors of cell substance interactions of drugs having known mechanisms of action with cells. Such descriptors can be compared using a plurality of techniques, such as a technique of creating "phylogenetic trees" of a statistical similarity between the descriptors from various drugs. In a present embodiment, scalar, numeric values can be converted into a nucleotide or amino acid letter. Once converted into a corresponding nucleotide representation, the descriptors can be analyzed and compared using software and algorithms known in the art for genetic and peptide sequence comparisons, such as GCG, a product of Genetics Computer Group, with company headquarters in Madison WI. In an alternative embodiment, numeric values for the fingerprints can be used by comparison techniques. A phylogenetic tree can be created that illustrates a statistical significance of the similarity between descriptors for the drugs in the database. Because the drugs used to build the initial database are of known mechanism, it can be determined whether a particular scalar value in a descriptor is statistically predictive. Finally, a compound fingerprint with no known mechanism of action can be queried against the database and be statistically compared and classified among the drugs in the database that the compound most resembles.

In a particular embodiment, relationships between measured morphometric properties and features of images and physiological conditions can be determined. Relationships can include, for example, treatment of different cell lines with chemical compounds, or comparing cells from a patient with control cells, and the like. In a presently preferable embodiment, a clustering can be performed on acquired image descriptors. Some embodiments can comprise statistical and neural

network - based approaches to perform clustering and comparisons of various descriptors. The foregoing is provided as merely an example, and is not intended to limit the scope of the present invention. Other techniques can be included for different types of data. In some embodiments, clustering and comparing can be performed on features extracted from cell images. In a presently preferable embodiment, procedures for comparisons and phylogenetic analysis of biological sequences can be applied to data obtained from imaging cells.

Select embodiments comprising such approaches enable the use of a broad array of sophisticated algorithms to compare, analyze, and cluster gene and protein sequences. Many programs performing this task are known to those of ordinary skill in the art, such as for example, the program Phylip, available at <http://evolution.genetics.washington.edu/phylip.html>, and other packages listed at <http://evolution.genetics.washington.edu/phylip/software.html>. However, select embodiments according to the present invention can comprise a technique of statistical classification, statistical clustering, distance based clustering, linear and non-linear regression analysis, self-organizing networks, and rule-based classification.

Embodiments can perform such analysis based upon factors such as numerical value, statistical properties, relationships with other values, and the like. In a particular embodiment, numbers in a numerical descriptor can be substituted by one or more of nucleic acid or amino acid codes. Resulting "pseudo-sequences" can be subjected to analysis by a sequence comparison and clustering program.

Other types of databases can also be provided according to other embodiments. The database includes details about the properties of a plurality of standard drugs. When the descriptor of a test compound is compared to the database, predictions about the properties of the test compound can be made using any known property of the other compounds in the database. For example, properties about a compound in the database could include structure, mechanism of action, clinical side effects, toxicity, specificity, gene expression, affinity, pharmacokinetics, and the like. The descriptor of a compound of unknown structure from a natural products library could be compared to the descriptors of compounds with known structure and the structure could be deduced from such a comparison. Similarly, such information could lead to better approaches to drug discovery research including target validation

and compound analogizing, as well as pre-clinical animal modeling, clinical trial design, side effects, dose escalation, patient population and the like.

According to the present invention, databases can be integrated with and complementary to existing genomic databases. Differential genomic expression strategies can be used for drug discovery using database technology. In one particular embodiment, cell data and cellular response data can be associated with a genetic expression profile assay to form a single assay. Live cells expressing fluorescence markers can be treated with a drug, imaged and analyzed for morphometry; and then analyzed for mRNA for expression. Such embodiments can provide rapid development of tools to link cellular behavior with functional genomics.

Database methods according to the present invention can be used to predict gene function and to assist in target validation. Databases that include genetic diversity, i.e., having cellular descriptors from cells of differing genetic backgrounds (tumor, tissue specific, and gene knock out cell lines), can provide the capability to compare cells of unknown genetic background to those in the database. Similarly, the descriptor of an unknown cellular portion in the presence of multiple drugs can be queried against the descriptors of the known markers in the database. For example, if an unknown gene is tagged with Green Fluorescent Protein (GFP), the database may be used to identify the cellular portions for which that unknown gene encodes.

According to the present invention, target validation and specialized cell-based assay screening can be performed using database systems and methods to serve as a universal high-throughput cell-based assay that can evaluate the molecular mechanism of drug action. As new genes are isolated and identified, a large collection of available gene-based knowledge is becoming available. From this large collection of new genes, potential protein targets can be identified using the genomic tools of sequence analysis and expression profiling. However, unless a gene mutation is tightly linked to a disease state, further validation of individual targets is a time consuming process, becoming a bottleneck in drug discovery. Furthermore, robotics and miniaturization are making "High Throughput Screening (HTS)" the industry standard, substantially reducing the time and cost of running a target-based biochemical assay. Therefore, it is now possible to routinely screen large libraries and use a resulting "hit" to validate the target. In such approaches, a specialized cell-based assay would be developed to test hits for each target. Since this often involves

the creation of cell lines expressing new markers, this stage may also become a bottleneck that cannot keep pace with HTS. In addition, these cell-based assays may not be amenable to high-throughput screening, making it difficult to test the increasing number of analogs arising from combinatorial chemistry.

5 In a particular embodiment according to the invention, a rapid characterization of large compound libraries for potential use as pharmaceutical products can be provided by predicting properties of compounds that relate to the compounds' potential as bioactive drugs. In many drug discovery situations, virtually millions of compounds can be passed through a HTS assay against a small number of
10 validated targets. These assays produce hundreds to thousands of potential hits. These hits can then be subsequently screened by a pipeline of secondary and tertiary screens to further characterize their specificity, often time completely missing non-specific interactions with other proteins. Techniques according to the present invention can provide a replacement to such screening operations by providing
15 information about cellular accessibility and mechanism of action for the hits coming from a HTS system. Furthermore, it can replace the biochemical HTS assay and allow rapid and accurate identification of attractive compounds from large libraries without an intervening biochemical assay. The cell information can be predictive of whether to continue into an animal model for each compound, and which animal model to
20 pursue.

 The principles of the present specifically contemplate a wide variety of research methodologies, or usage scenarios, implementing these principles. The following discussion of three such scenarios is by way of illustration and not limitation. Study of the principles enumerated herein will render evident to those
25 skilled in the art certain additional methodologies or usage scenarios enabled by the teachings hereof. The present invention specifically contemplates all such modifications. The following description presents some specific embodiments and scenarios that represent a broader use of cellular phenotypic data and characterizations to deduce mechanisms of action and other features of cellular
30 responses to various stimuli. Such procedures generally involve producing a quantitative cellular phenotype based upon two or more cellular attributes and then comparing that phenotype to phenotypes previously stored and indexed. Such

procedures make use of databases or other repositories of biological information. The invention is not limited to the specific embodiments described here.

Considering first the procedure 2000 depicted in Figure 20, a compound has been identified as having a particular cellular activity. See 2004. For example, a compound may be found to inhibit the growth of certain cancer cell *in vitro* by a specific and desired mechanism of action. This may be a particular company's "gold standard."

Next, the compound is analyzed at 2006 in terms of its effect on one or more cell lines. More specifically, the compound is linked, virtually, to a particular phenotype. Two or more values or measures of cellular attributes characterize that phenotype. These attributes are quantified in the context of specific cellular markers.

In one example, the cellular marker is an organelle such as a nucleus or Golgi apparatus. Measured attributes useful for characterizing an associated phenotype include geometric parameters (e.g., size, shape, and/or location of the organelle) and composition (e.g., concentration of particular biomolecules within the organelle).

The phenotype may be characterized by administering the compound of interest to various cell lines and in various concentrations. In each example within this matrix, the attributes of interest are measured. Ultimately, certain phenotypic features (combinations of attribute values) are associated with the compound of interest. These features provide a template for the phenotype.

Next, using the phenotype as identified at 2006, the process identifies other compounds providing similar features. The goal here is to present a list of compounds having a mechanism of action similar to that of the compound that started the process. This allows researchers to identify a mechanism of action, if not already known, for their compound and to draw conclusions based upon their compound's link to other known compounds (which may not be chemically/structurally similar to the compound of interest).

Identifying similar compounds based upon phenotype can take many paths. Most will involve some mathematical basis. For example, the phenotype defined at 2006 can be represented as a fingerprint or vector comprised of multiple scalar values of cellular attributes (as described above). The phenotype representation can then be compared against known phenotypes characterized by the same format

(e.g., they are all characterized as vectors having the same attribute set, but with different values of the attributes). The comparison may be as simple as a Euclidean distance or more sophisticated as a neural network or multivariate statistical correlation.

5 The known compounds and associated phenotypes may be stored as database records or other data structures that can be queried or otherwise accessed as part of the identification procedure. The compounds may also be associated with other relevant data such as clinical toxicity, cellular toxicity, hypersensitivity, mechanism of action, etc. (when available).

10 Compounds found to be sufficiently similar to the starting compound are returned for consideration by researchers. A data processing system may rank such compounds based on degree of similarity to the starting compound. In some cases, the system may even provide similarity scores associated with the listed compounds.

15 Often researchers wish to determine whether their particular compound has clinical or biochemical effects beyond those that they are already aware of. In a typical scenario, the compound of interest was selected based upon its strong binding a target or its stimulation or inhibition of cell growth in a particular cell line. The process associated with 2010 has likely identified the compound of interest as having
20 a particular mechanism of action based on phenotypic similarity to other compounds having a similar mechanism of action. However, within the region of biochemical space, there may be subspaces (characterized by subphenotypes) that correspond to separate properties. For example, within the phenotypic space associated with one mechanism of action, there may be subspaces associated with clinical toxicity,
25 cellular toxicity (likely overlapping the clinical toxicity space), and little or no toxicity. Obviously, a researcher would like to know whether her compound is likely to be toxic.

 Thus, the process 2000 may include characterizing the compound of interest in terms of its distance from (i.e., similarity to) specific phenotypes having
30 known characteristics. In a typical example, the known characteristic is toxicity. This feature allows the researcher to quantify her compound in terms of mechanism of action AND toxicity (or in terms of two or more other relevant properties associated

with phenotype). To allow simple ranking or characterization, compounds of interest may be scored according to a simple or weighted Boolean expression.

A second scenario of interest is depicted in Figure 21. This scenario again defines a phenotype in terms of a quantifiable vector or other measure.

5 However, rather than using a compound of interest to generate the phenotype, some other cellular stimulus is used to generate the phenotype.

As shown, a process 2100 begins with receipt of cells of interest. See 2104. In many situations, the cells are produced by a genetic or epigenetic process that affects the expression level or activity of a particular protein. More generally,
10 any cellular stimulus (e.g., radiation level and type, gravity level, magnetic field, acoustic perturbations, etc.) can be used to generate the cell line of interest. Importantly, this stimulus affects the phenotype and can be correlated therewith.

In the context of drug discovery, a gene encoding for a particular target can be genetically knocked out, underexpressed, overexpressed, expressed in a non-
15 native state, etc. This may be accomplished via standard procedures involving genomic modification, translation or transcription apparatus modification (e.g., use of antisense nucleic acids), blocking target activity (using antibodies to a receptor site for example), and the like. These processes will generally affect the phenotype in some quantifiable way. Importantly, they clearly and unambiguously define a cellular
20 phenotype associated with altering the activity of the target protein.

At 2106, the process involves measuring one or more cellular features from the cell line of interest to define/quantify the phenotype. This may be accomplished as described above with reference to 2006. Next, at 2108, the cellular phenotype generated in this manner is used to identify and rank a set of compounds
25 associated with the phenotype. This operation may proceed in the manner of operations 2008 and/or 2010 from Figure 20.

Finally, at 2110, the process clusters the compounds returned at 2108 by a mechanism of action. The operation 2106 has tightly bound a mechanism of action to a phenotype. Various compounds characterized and stored in a system
30 database may be tentatively assigned a mechanism of action or may have no suggested mechanism of action. By matching their virtual phenotype to the phenotype generated at 2106, one can create or strengthen an association between the compounds and mechanism of action relevant to the stimulus at 2104.

Considering now Figure 22, a third scenario is depicted. This scenario again involves using a virtual phenotype to glean information relevant to a mechanism of action or other cellular activity. In this case, assay data from a group of compounds (e.g., a primary or focused library) is used to elucidate a phenotype.

5 As shown, a process 2200 begins by identifying a target protein. See 2204. Then, at 2206, the process involves identifying positive and negative biochemical hits. More generally, this may involve ranking a number of compounds based upon their interaction with the target. In a specific case, the compounds are ranked based upon their binding affinities to or ability to inhibit the enzymatic activity
10 of the target protein.

After the compounds have been characterized in some manner based upon their interaction with the target, they are used to define a cellular phenotype. See 2208. Generally, the techniques to accomplish are the same as described with reference to operation 2006 of Figure 20. In this case however, one may obtain a
15 strong correlation between mechanism of action (involving the target) and phenotype by using multiple of the compounds identified at 2206. For example, some of the "best hits" may be administered to cell lines in various concentrations. And some of the least effective compounds may also be administered. Cellular attributes that are more strongly exhibited with increasing concentration of the best hits (and not
20 exhibited or exhibited only weakly upon administration of the negative hits) can be used to define the virtual phenotype. In a related approach, compounds having widely varying levels interaction with the target are administered to cells. Those cellular attributes that vary linearly or at least monotonically with the degree of interaction between the target and compound represent attributes that can be used to define the
25 virtual phenotype.

After the cellular phenotype has been defined, previously characterized compounds may be clustered with that phenotype. See 2210. As with operation 2110 of Figure 2, this may create or strengthen an association between a mechanism of action and various compounds in a database.

30 Finally, and optionally, procedure 2200 may provide a "higher resolution" mechanism of action for the compounds identified at 2206. See 2212. Presumably interaction with the target suggests a specific mechanism of action or at least some aspect of a mechanism of action. However, a given target may participate

in a larger cellular mechanism of action – unknown to researchers. Further, a compound may that binds with the target may participate in multiple mechanisms of action – some of which do not involve the target. By linking the target (and its positive hits) to a particular phenotype, some of these additional cellular level activities can be elucidated. The defined phenotype may have been previously identified as associated with other mechanisms of action or higher resolution mechanisms of action. Thus, the phenotype identified at 2208 can be leveraged to generate a higher resolution mechanism of action at 2212.

As suggested in the above discussion, compounds and associated phenotypes may be stored as database records. Such databases can take on many flavors. In one example, a database includes various pieces of information relevant to oncology. Such database may include numerous compounds classified by cellular phenotype, mechanism of action, toxicity, etc. More specifically, the database may include data on commercially available compounds clustered by cellular phenotypes corresponding to mechanisms of action. Further the databases of interest may extended or combined (via standard relational tables and algebra for example) to include additional data such as pharmacology data, cellular genomics data, gene expression data, protein expression data, etc. In a specific example, the database includes measurements made on a subset of the NCI60 cell lines, using DNA, Golgi apparatus, and/or microtubules as markers for defining the phenotypes. Other data includes dosage response information, variation in effect over time, etc. The compounds populating the database could include known National Cancer Institute oncology study compounds. In a specific embodiment, the compound set includes some or all of the compounds mentioned in the article “A gene expression database for the molecular pharmacology of cancer,” Nature Genetics, 24, pp. 236-244 (March 2000).

Various biological analyses may be conducted to develop additional information for characterizing compound mechanisms of action, etc. For example, a cell count analysis may be used to develop dose response curves, GI 50 data, etc. The cell cycle may also be analyzed to find out how various stages in the cycle vary in response to particular stimuli. The Golgi apparatus may be analyzed to determine whether it is in a normal state, a dispersed state, a diffused state, etc. As another example, tubulin may be analyzed to determine whether it is normal, de-polymerized,

over-polymerized, bundled, etc. Obviously, combinations of such analyses may be performed. For example, properties of the Golgi apparatus or tubulin may be analyzed over one or more cell cycles.

5 In some embodiments, techniques according to the present invention can provide tools for the later stages of drug development such as clinical trial design and patient management. The properties of known drugs, such as clinical trial and patient response information, will be used in a similar fashion as the pre-clinical information to provide predictions about the properties of novel compounds. Because the human cell is the locus of drug action, a database containing drug-cell interactions
10 will be able to provide predictive value for this aspect of drug development.

Although the above has generally been described in terms of specific hardware, software, and methods, it is understood that many alternatives can exist. In particular, the present invention is not limited to a particular kind of data about a cell, but can be applied to virtually any cellular data where an understanding about the
15 workings of the cell is desired. Thus, in some embodiments, the techniques of the present invention could provide information about many different types or groups of cells, substances, and genetic processes of all kinds. Of course, one of ordinary skill in the art would recognize other variations, modifications, and alternatives. Some examples according to the present invention are provided below.

20

EXPERIMENTS

To prove the principle and demonstrate the objects of the present invention, experiments have been performed to determine the effects of manipulations on cell structure using imaging and analysis techniques applied to a variety of
25 situations. These experiments were performed by growing multiple cell lines in the presence of multiple compounds, or substances. Cells were fixed and stained with fluorescent antibodies or labels to multiple cellular portions. One or more images of the cells were then obtained using a digital camera. Descriptors were built by quantifying and/or qualifying patterns of one or more feature from each image in the
30 cell lines under study. A database was built from the descriptors. As the database grows, it should be able to predict the mechanism of action of an unknown drug by comparing its effect with the effects of known compounds or to identify data clusters within large libraries of compounds.

In a first experiment, an automated method to count the number of cells and differentiate normal, mitotic, and apoptotic cells was created. Approximately, 5,000 HeLa cells were plated per well in a 96 well plate and grown for 3.5 days. The cells were fixed with -20° MEOH for 5 minutes, washed with TBS for 15 minutes, and then incubated in 5 mg/ml Hoechst 33342 in TBS for 15 minutes. Then, 72 images were collected with a 40x objective and 75 ms exposure time.

The analysis was performed on objects that met a certain size criteria that was based on 1) measuring the size of objects in the image that were clearly not cells and 2) excluding the first peak of the area histogram (Fig. 8B values 1-4654).

Histograms of the individual object data were generated for each type of feature. Fig. 8A shows the histogram for average intensity, and Fig. 8B shows histogram data for the area of each object. Fig. 8C shows the scatter plot of the average intensity vs. the area of all of the objects. The pattern of the scatter plot showed an interesting pattern: a large cluster of cells in one region of the graph, with a scattering of object points in other regions. Because mitotic structures are identified as particularly bright objects, most likely due to the biological fact that the chromatin is condensed, the original Hoechst images could be used to identify which cells were either undergoing mitosis, or otherwise looked abnormal. Manual inspection of 917 cells resulted in the classification of each object. Fig. 8D shows a graph where each type of cellular classification is delimited. This graph clearly shows that the mitotic nuclei are brighter than the interphase nuclei. Further, the different phases of the cell cycle can be separated using these two features. Figs. 8E-8F show bar graphs of the average and standard deviations of the areas and average intensities for each cell classification type. These graphs show that interphase nuclei are statistically less bright than mitotic nuclei and that telophase nuclei are statistically smaller than other mitotic nuclei.

Each image was thresholded to an intensity level of 20. A standard area value was set at 9500 pixels. Automated information gathering about all of the objects was done and collected into an Excel spreadsheet (for more information see, section on imaging system). The following information was recorded:

IMAGE NAME
OBJECT #

AREA
STANDARD AREA COUNT
PERIMETER
FIBER LENGTH
FIBER BREADTH
SHAPE FACTOR
ELL. FORM FACTOR
INNER RADIUS
OUTER RADIUS
MEAN RADIUS
AVERAGE INTENSITY
TOTAL INTENSITY
OPTICAL DENSITY
RADIAL DISPERSION
TEXTURE DIFFERENCE MOMENT
EFA HARMONIC 2, SEMI-MAJOR AXIS
EFA HARMONIC 2, SEMI-MINOR AXIS
EFA HARMONIC 2, SEMI-MAJOR AXIS
ANGLE
EFA HARMONIC 2, ELLIPSE AREA
EFA HARMONIC 2, AXIAL RATIO
EFA HARMONIC 3, SEMI-MINOR AXIS

The following results were obtained:

- 1,250 objects were counted
- 201 of those objects has standard area counts > 2 (area > 19000 pixels)
- 195 objects had areas < 6000 pixels
- 1529 objects estimated in total
- 1328 object areas are > 6000 pixels
- The data was reduced to 917 objects that were $6000 < \text{area} < 19000$
- For the 917 objects a scatter plot of area vs. average intensity and a histogram of the average intensity were generated.

- 116 objects that had average intensity intensities > 60 were manually looked at to determine their morphology.

- Of those 116 objects:

6 were dead or indistinguishable

5

4 were interphase

30 were prophase

32 were metaphase

24 were anaphase

20 were telophase (10 pairs)

10

- 12 prophase objects were missed because of gray scale cut off. (8 of those prophase cells had gray scale values > 57 , as did 7 interphase)
- 1 telophase object was missed because it was too small (< 6000)
- 1 prophase object was missed because it was too big (> 1900)
- 16 mitotic objects were missed because they were parts of objects with standard count > 2 .

15

In sum, out of 917 single objects, the analysis correctly identified 106 out of 130 mitotic objects, or (81% predictive, 91% of identified mitotics). Out of 917 single objects, the analysis incorrectly identified only 10 non-mitotics as mitotics (1% total, 8% of identified mitotics); 14 mitotics as interphase (1.4% total, 1% interphase). An automated classification system that would automatically assign values to each object using these or other measurement features can thus be developed, utilizing the principles set forth herein.

In a second experiment, the effects of Taxol on MDCK cells and the different types of morphological effects were observed. A plurality of MDCK cells grown in 96 well plates were treated with Taxol for 4.5 hours at different concentrations (10 uM-1pM). They were then fixed, labeled with Hoechst, and imaged.

This experiment used a labeling protocol comprising: MEOH fix at – 20°, Wash in PBS, Block in PBS/BSA/Serum/Triton-X 100, Incubate with 5 µg/ml Hoechst 10 minutes, and wash.

Cells were inspected for different morphologies and manually counted at each different drug concentration in one well. Fig. 9 shows example images from each drug concentration and the different types of morphologies and cells are highlighted. Fig. 10 shows the distribution of each morphology within the cell population as a function of drug concentration. The higher the concentration of Taxol, the larger proportion of cells underwent apoptosis, and the fewer number of normal mitotic cells were detected.

In a third experiment, the purpose was to determine whether the automated analysis methods developed in the first experiment can detect differences in Hoechst morphology in the presence of 6 known compounds at one concentration and exposure time in one cell line. In this experiment, HeLa cells were separately treated with 6 compounds with known mechanism of action. The quantitative methods described in the first experiment were applied to the Hoechst images.

Approximately 5,000 HeLa cells per well were plated in a Costar black-walled 96 well tissue culture treated plate and left to recover in the incubator for 24 hours. After this time, 10 ug/mL of cytochalasin D (CD), Taxol, hydroxyurea, vinblastine, nocodazole, and staurosporine was added to different wells at a 1:100 addition in DMSO.

The cells were incubated in the presence of drug for 24 more hours. After 24 hours, the cells were removed and fixed as in the first experiment. Then, 9 images per well were collected of the Hoechst staining using a 10x objective.

The low magnification images taken of Hoechst were run through the automated image analysis method described in the first experiment. Plots of the average intensity and area were made of each compound. Fig. 11 shows the scatter plots of the compounds. The scatter plots of each compound are visually distinct. For example, cells treated with CD are smaller than control, and cells treated with Hydroxyurea are larger and brighter. Furthermore, the number of cells per well was very different (data not shown).

The effects of different compounds can be clearly and automatically distinguished by identifying changes in cellular morphology. This method can also be used to count adherent cells.

The next experiment was to develop clustering algorithms that assign statistically meaningful values to the representative two dimensional data shown in Fig. 10, and even more complicated clustering of all of the multidimensional data that can be extracted across one, and multiple images.

A fourth experiment was performed to obtain high magnification images of two markers in the presence of drugs. In this experiment, HeLa cells were treated with 80 generic compounds with known mechanism of action. The quantitative methods described in the first experiment were applied to the Hoechst images.

Approximately 5,000 HeLa cells per well were plated in a Costar black walled 96 well tissue culture-treated plate and left to recover in the incubator for 24 hours. After this time, 10 ug/mL of each compound from the Killer Plate from Microsource Discovery Systems (Gaylordsville, CT) was added to different wells at a 1:100 addition in DMSO. The cells were incubated in the presence of drug for 24 more hours. After 24 hours, the cells were removed and fixed as in the first experiment. In addition to being labeled with Hoechst 33342 (against chromatin), cells were also labeled with 1 unit of rhodamine-conjugated phalloidin (against actin) for 30 minutes.

The 96 well plate was imaged twice. Once, 9 images per well were collected of the Hoechst staining using a 10x objective. After this, one image per well of both the phalloidin and Hoechst staining was collected using a 40x objective.

The resulting high magnification images were analyzed qualitatively and distinct pattern differences were detected in both the Hoechst and phalloidin images. Fig. 12 shows three example images from the experiment. The top row is the Hoechst staining, and the bottom row is the phalloidin staining from the same well. The columns show the images from wells treated with just DMSO (control), cytochalasin D, and Colchicine. The morphology of each marker is different in the presence of each drug. Interestingly, there is an effect in the morphology of the chromatin in the Hoechst image of cytochalasin D, which directly targets the actin cytoskeleton (and thus there is an expected effect in the phalloidin image). Also, there is an effect on the actin cytoskeleton, compared to control, in the presence of colchicine that directly targets the microtubule network.

The low magnification images were analyzed as described in the first experiment, and different patterns were seen in both the average intensity vs. area plots, and in the number of cells per well (data not shown). Thus, changes in patterns of a marker that is "down-stream" from the direct target of a compound are detectable. Automated image analysis protocols for actin and other markers can be developed similarly, again utilizing the principles set forth herein.

A fifth experiment was performed to test quadruple labeling of 9 different cell lines grown in normal conditions. In this experiment, NCI-H460, A549, MDA-MD-231, MCF-7, SK-OV-3, OVCAR-3, A498, U-2 OS, and HeLa cells were plated. Then, the cells were fixed and stained for portions of the each cell known as DNA, tubulin, actin, and Golgi.

The following table summarizes the procedures for this experiment:

Action	Active Ingredient/Notes	Buffer	Vol/ well	Desired Time	Temp
Remove media	NOTE: gently by pipetting, not aspiration				
Fix	4% Formaldehyde	PBS	100µl	20 min	rt
Wash		TBS	100µl	5 min	rt
Wash		TBS	100µl	5 min	rt

Permeablize	0.1% Triton X-100	TBS	100 μ l	10 min	rt
Permeablize	0.1% Triton X-100	TBS	100 μ l	10 min	rt
Block	% BSA % Serum Filter sterilize before use	TBS w/azide	100 μ l	1hr or o/n	rt or 4°C
Primary Antibody	1:1000 dilution of DM1 α	TBS + 1% BSA + 0.1% TX-100	50 μ l	1hr or o/n	rt or 4°C
Wash		TBS	100 μ l	5 min	rt
Wash		TBS	100 μ l	5 min	rt
Wash		TBS	100 μ l	5 min	rt
Fluorescent Stain	FITC lens culinaris 1:500 Rhodamine-Phalloidin 1:500 CY5 goat anti-mouse 1:100	TBS + 1% BSA + 0.1% TX-100	50 μ l	1 hr.	rt, dark
Wash		PBS	100 μ l	5 min	rt, dark
Hoechst	1:1000 dilution of 5mg/ml	TBS	100 μ l	15 min	rt, dark
Wash		PBS	100 μ l	5 min	rt, dark
Wash		PBS	100 μ l	5 min	rt, dark
Wash		PBS	100 μ l	5 min	rt, dark
Store		PBS	200 μ l	1 month	4°C

Cells were plated out at different densities for 48 hours. Cells were fixed and labeled by the above method. Cells were imaged using an automated imaging system that collected 9 images from each marker using a 10x objective.

Higher magnification images were collected of a few cells for demonstration purposes.

In this experiment, each cell line demonstrated different morphological patterns as determined by phase. For example, A549 cells are much more compacted than OVCAR-3 cells as determined by phase contract imaging (data not shown). The different fluorescent markers showed even bigger differences between different cell lines. Figs. 13 and 14 show 4 panels of each marker for A549 (Fig. 13) and OVCAR-3 cells (Fig. 14). The markers are Hoechst (upper left), Phalloidin (upper right), Lens culinaris (lower left), and DM1a antibody (lower right). The following table summarizes the qualitative differences between these images:

MARKER	A549	OVCAR3
Hoechst/DNA	small	large
Phalloidin/actin	fuzzy	crisp - many stress fibers
Lens culinaris/Golgi	compact	Disperse/punctate
DM1alpha Tubulin	perinuclear	evenly distributed

Higher magnification images were taken of the OVCAR3 cells. Fig. 15 shows the same markers at 20x, and Fig. 16 shows the markers at 40x. While the highest magnification images show the most detail, these images illustrate that very little morphological or feature information is lost in the 10x images.

These data exemplify the differences in morphology seen between different cell types. Thus the automated image analysis software can be customized for each marker in each cell type. Different drugs should effect these morphologies differentially.

An automated quantification method for each marker and cell line can be similarly developed.

A sixth experiment was conducted with a more sophisticated software package and to develop more flexible image recognition algorithms. In this experiment, prototype image features extraction was performed using MatLab programming language with image toolbox and SDC morphology toolboxes. Algorithms are being developed that will automatically identify objects on images and

to measure various morphological and feature parameters of these objects. Many different features for each of the cellular markers were acquired.

An example of a MatLab program called "AnalyseDNA" that takes as an input an unlimited number of images, identifies individual objects in these images based on either their intensities, or based on edge-detection algorithms, and extracts a number of morphological and intensity characteristics of these objects. A copy of this program follows:

Listing of the AnalyseDNA.m program and of some of the
supporting subroutines

```

10 function files_analysed = AnalyseDNA(filemask, outpath,
    nx, ny, filter_range, dext, modifier, sfname)
% AnalyseDNA performs measurements on files of DNA images
% V1. EV 2-11-99; 2-15-99; 2-16-99
15 %
% files_analysed = AnalyseDNA(filemask, outpath, nx, ny,
    filter_range, dext, modifier, sfname)
%
% PARAMETERS:
20 %   ALL PARAMETERS ARE OPTIONAL
%
%   FILEMASK - mask for file names to be analyzed
    INCLUDING PATH(for example c:\images\*.tif)
%   DEFAULT '.\*.tif' (all *.tif files in the current
25 directory).
%
%   OUTPATH - path to a directory where all the output
    files will be placed.
%   DEFAULT - output is saved in the same directory
30 which contains images
%
%   NX, NY - number of individual images in montage
    images along X and Y axes (DEFAULT 1)
%
```

```
%    FILTER_RANGE - 3 col-wide array (or[]). Specifies
how data is filtered when summary is calculated
%    this parameter internally is passed to GetDNADData
and then to GetSummaryData - see these
5 %    functions for details. For example: [2 2 Inf; 6 100
8000] will case all raws of data for which
%    values in column 2 are less than 2 and all raws
where values in column 6 are less than 100 or
%    more than 8000 to be excluded from all
10 calculations of a summary.
%    DEFAULT - [] (means do not filter, summarize all
data)
%
%    DEXT - string. Extension for data files being saved.
15 %    DEFAULT 'dat';
%
%    MODIFIER - this modifier is 'SUMMARY', summary file
is created;
%    'SUMMARY ONLY' - only summary is generated,
20 data for individual files are not saved
%
%    sfname - string. File name of a summary file
%    DEFAULT 'summary[date].dat'
%
25 % OUTPUT:
%
%    AnalyseDNA works on image files or montages. For
each image file it creates a tab-delimits file of
measured
30 %    parameters of all the objects in the montage with
the same base name as a montage file and extension
specified
```

```
%    by dext parameter (or .dat by default) and file
'errors[date].err' - with the list of files that matched
the
%    filemask but could not be processed.
5 %    If 'summary' or 'summary only' modifier is
specified, it also creates a single file
'summary[date].dat' (or
%    different extension, if specified by DEXT) which
contains summary information for all analyzed files.
10 %
%    ALL OUTPUT FILES are saved in a directory specified
by OUTPATH parameter
%
%    RETURNS *files_analysed* - number of files that have
15 been successfully processed.
%
%    Column designations in the output files are
described in GetDNADData
%
20 % FILE NAME CONVENTIONS
%    AnalysedDNA attempts to identify a number for each
file to identify the file in summary output.
%    It does that by looking for the first space or
underscore, followed by a number and then takes
25 %    as many successive numbers as it can find. If it
fails to identify a number it assigns a
%    default which is -1
%
%
30 % SEE ALSO GetDNADData, GetSummaryData
%
% TO DO    improve error handling in opening and writing
files (GLOBAL error_file ?)
```

```
%          include procedures for writing text headers
into the output files

.

if nargin > 8
5   error ('Wrong number of input parameters');
end
if nargsout >1
    error ('Wrong number of output parameters: only one
allowed');
10  end

% set defaults
need_summary = 0;
summary_only = 0;
15  use_default_outpath = 0;
datestring = datestr(floor(now));
if nargin == 7      % set default summary file name
    sfname = ['summary' deblank(datestring)]; % extension
will be appended later based on dext
20    if deblank(upper(modifier)) == 'SUMMARY'
        need_summary = 1;
    elseif deblank(upper(modifier)) == 'SUMMARY ONLY'
        need_summary = 1;
        summary_only = 1;
25    else
        error (['Wrong parameter: unknown modifier '
modifier]);
    end
end
30

if nargin == 5
    % default data file extension
    set_dext = 'dat';
end
```

```
    if nargin == 4
        % default filter range
        filter_range = [];
    end
5   if nargin == 3
        ny = 1; % default number of images in montage along Y
    end
    if nargin == 2
        nx = 1;
10  end
    if nargin == 1
        use_default_outpath = 1;
    end
    if nargin == 0
15  filemask = '*.tif'
    end

    % check parameters
    if ( ~ischar(filemask) | ~ischar(dext) | ~ischar(sfname)
20  )
        error('Wrong parameter type: filename, filepath,
dext and sfname should be strings');
    end
    if ( ( size(nx) ~= [1 1] ) | ( size(ny) ~= [1 1] ) )
25  error ('Wrong parameter type: nx and ny should be
scalars (1x1 arrays)');
    end
    if (~isempty(filter_range) & size(filter_range, 2) ~= 3)
        error ('Wrong parameter type: filter range should be
30  [] or 3 - cols-wide array');
    end
    % end testing parameters

    % Generate list of files to process
```

```
datapath = getpath(filemask);
if use_default_outpath == 1
    outpath = datapath;
5  end
if exist(outpath, 'dir') ~= 7
    error(['Path ' outpath, 'not found. Exiting..']);
elseif exist(datapath, 'dir') ~= 7
    error(['Path ' datapath, 'not found. Exiting..']);
10 end

sfname = makefullname(outpath, sfname, dext);
if need_summary == 1
    if exist(sfname, 'file')
15     disp(['File ', sfname, 'already exists!']);
        input ('Press ^C to abort, Enter to delete and
continue');
        delete(sfname);
    end
20 end

flist = FileList(getfname(filemask), datapath);
numfiles = size(flist, 1); % total number of files to
25 process
disp(['About to process ', num2str(numfiles), ' files']);
%DEBUG - commented out "input" to run from Wrod
input('Press ^C to abort, Enter to continue');

30 % main loop where the job gets done:
error_file = makefullname(outpath, ['error' datestring
'.err']);
num_processed = 0;
num_error = 0;
```

```
for i = 1:numfiles
    % first generate file name for a data output file
    current_fullname = flist(i, :); % full name with path
    and extension
5    current_datafile = makefullname(outpath,
    makefname(getbasefname(current_fullname), dext) );

    %extract number from a filename
    fnumber = getfilenumber(current_fullname);
10
    % load an imagefile, record errors
    read_error = 0;
    try
        I = imread(current_fullname);
15        %DEBUG
        disp(['Image file #', num2str(fnumber), '
loaded']);
    catch
        % record file-opening error in an error_file
20        read_error = 1;
        num_error = num_error +1;
        msg = [current_fullname ': ' lasterr];
        add_error_msg(error_file, msg);
    end
25
    % extract and write data to a file in outpath
    if read_error ~=1
        if (need_summary == 0)
            %DEBUG
30            disp(['Starting analysis of file #',
num2str(fnumber), '.']);
            current_data = GetDNADData(I, nx, ny, fnumber);
            %DEBUG
```

```

        disp (['Finished analysis of file #',
num2str(fnumber), '.']);
        %load current_data.mat 'current_data';
        write_data(current_data, current_datafile);
5         else      %summary needed
            %DEBUG
            [current_data, current_summary] = GetDNAData(I,
nx, ny, fnumber, filter_range);
            %load current_data.mat 'current_data';
10         %load current_summary.mat 'current_summary';
            write_summary (current_summary, sfname);
            if summary_only ~= 1
                write_data(current_data, current_datafile);
            end
15         end
        end
    end % of the main for loop
    num_processed = numfiles - num_error;

20  %=====end function AnalyseDNA()
    %=====

    %=====
    %=====

25  function result = add_error_msg(filename, msg)
    % adds string MSG to an errorfile FILENAME
    % returns 1 if success, 0 if failure

    err_FID = fopen(filename, 'at');
30  if err_FID == -1
        warning(['Can not open error file ' filename]);
    else
        fprintf(err_FID, '%s\n', msg);
        fclose(err_FID);
    end

```



```

end
%=====end function add_error_masg()
=====

5 %=====
=====
function N = getfilenumber(fname)
% returns the first number extracted from a file name
(string) or -1 if fails to extract any number
10 numbers = NumbersFromString( getfname(fname) ); % vector
of all numbers encoded in the name

                                % (but not in the path, even if
present)
15 if isempty(numbers)
    N = (-1);    % return -1 if no numbers found in the
name
else
    N = numbers(1);
20 end

%===== end function getfilenumber()
=====

25 %=====
=====
function result = write_data(data_array, file_name)
% writes data in a data_array in a tab-delimited ascii
file.
30 % result is 0 if success and -1 if failure
% if file_name exists, overwrites it
result = -1;
try
    fid = fopen(file_name, 'wt');

```

```

        if fid ~= -1
            for k = 1:size(data_array, 1)
                fprintf(fid, '%g\t', data_array(k, :));
                fprintf (fid, '\n');
5         end
            test = fclose(fid);
            result = -1;
        catch
            result = -1;
10    end

%===== end function write_data()
=====

15 %=====
=====

function result = write_summary (s_vector, file_name)
% appends summary vector s_vector to a file_name (ASCII
tab-delimited file).
20 % if file_name does not exist, creates it.
% result is 0 if success and -1 if failure
%
result = -1;
try
25     % debug
        fid = fopen(file_name, 'at');
        result = fprintf(fid, '%g\t', s_vector);
        result = fprintf(fid, '\n');
        result = fclose(fid);
30     result = 0;
    catch
        result = -1;
    end
end

```

```

% ===== end function write_summary()
=====

function Data = GetObjectsData(I, Ilabel)
5 % GetObjectsData returns array measurements of objects in
  image "I" masked by "Ilabel"
  % EV 2-3-99; 2-10-99
  % OData = GetObjectsData(I, Ilabel) returns an array of
  morphological and intensity measurements
10 %   taken from a grayscale image "I". Objects are
  identified on a mask image Ilabel, usually
  %   created by bwlabel()
  % OUTPUT:
  % Each row in the output array OData represents
15 individual object
  % columns contain the following measurements:
  %
  %   1 - Index ("number" of an object);      8 -
  Solidity;
20 %   2 - X coordinate of the center of mass; 9 - Extent;
  %   3 - Y coordinate      "-"; 10 - Total
  Intensity;
  %   4 - Total Area (in pixels);              11 - Avg.
  Intensity;
25 %   5 - Ratio of MajorAxis/MinorAxis;        12 - Median
  Intensity;
  %   6 - Eccentricity;                        13 - Intensity of
  20% bright pixel
  %   7 - EquivDiameter;                       14 - Intensity of
30 80% bright pixel
  %
  % For details on morphological parameters see information
  on MatLab imfeature();

```

```
% Intensity parameters are either obvious or are
documented in comments in this file.
% Procedures in this file are documented in notebook file
"MATLAB Measuring Nuclei (1) 1-29-98.doc"

5
if (nargin ~= 2)
    error ('function requires exactly 2 parameters');
end
if (nargout ~= 1)
10    error ('function has 1 output argument (array X by
    14)');
end

% finished checking arguments

15
% first collect morphological parameters in a structure
array:
ImStats = imfeature(Ilabel, 'Area', 'Centroid',
    'MajorAxisLength',...
20    'MinorAxisLength', 'Eccentricity', 'EquivDiameter',
    ...
    'Solidity', 'Extent', 8 );

% now convert it into array (matrix) while collecting
25 intensity data for each object:

%preallocate output array:
numobjects = size(ImStats, 1);
OData = zeros(numobjects, 14);
30 %now convert ImStats into array and add intensity data to
it
for k=1:numobjects
    OData(k, 1) = k;
    OData(k, 2) = ImStats(k).Centroid(1);
```

```

        OData(k, 3) = ImStats(k).Centroid(2);
        OData(k, 4) = ImStats(k).Area;
        OData(k, 5) = (ImStats(k).MajorAxisLength) /
        (ImStats(k).MinorAxisLength);
5         OData(k, 6) = ImStats(k).Eccentricity ;
        OData(k, 7) = ImStats(k).EquivDiameter;
        OData(k, 8) = ImStats(k).Solidity;
        OData(k, 9) = ImStats(k).Extent;

10         % now collect and assign intensity parameters from
        image I

        object_pixels = find( Ilabel == k);
        object_area = size(object_pixels, 1); %same as total
15 number of pixels in the object
        object_intensities = double(I(object_pixels)); %
        need to convert to double to do math
        sorted_intensities = sort(object_intensities); %
        will need to get median, 20% and 80% pixels
20         total_intensity = sum(object_intensities, 1);
        avg_intensity = total_intensity / object_area;
        median_intensity = sorted_intensities( floor(
        object_area/2 ) + 1 );
        pix20 = sorted_intensities( floor(object_area*0.2)+1
25 ) ; %brightest pixel among dimmest 20%
        pix80 = sorted_intensities( floor(object_area*0.8)+1
        ) ;

        OData(k, 10) = total_intensity;
30         OData(k, 11) = avg_intensity;
        OData(k, 12) = median_intensity;
        OData(k, 13) = pix20; %brightest pixel among dimmest
        20%

```

```

        OData(k, 14) = pix80; %dimmiest pixel among brightest
    20%
    end %for

5   %===== end function
    GetObjectsData()=====

    function Imask = MaskDNA1(I);
10   % MaskDNA1 - generates binary mask for cell nuclei
    through edge detection
    % EV 1-22-99; 2-6-99; 2-10-99
    % Imask = MaskDNA1(I)
    % PARAMETERS
15   %   I - intensity image (grayscale)
    % OUTPUT
    %   Imask - BW image with objects from I
    %
    % For more details see Notebook Matlab_DNA_masking1_1-22-
20   99.doc
    % Uses SDC Morphology Toolbox V0.7

    if (nargin ~= 1)
        error('Wrong number of input parameters');
25   end
    if (nargout ~= 1)
        error('Wrong number of output parameters: one output
        argument should be provided');
    end
30

    Imask = edge(I, 'canny');
    Imask = mmdil(Imask, mmsecross(1));
    Imask = mmero ( mmc1ohole(Imask,mmsecross(1)));

```

```

Imask = mmedgeoff(Imask, mmsecross(1));
% note that mmedgeoff this command removed FILLED OBJECTS
but not touching OUTLINES.
% these outlines can be removed by filtering:
5 Imask = medfilt2(Imask, [5 5]);

%=====end MaskDNA1 =====

```

10 Given the list of image files or montages of images as an input, this program creates an individual file for each image that contains the following quantitative measurements for all objects identified in the image:

- | | |
|---|------------------------------------|
| 1 - Index ("number" of an object); | 8 - Solidity; |
| 2 - X coordinate of the center of mass; | 9 - Extent; |
| 15 3 - Y coordinate "-"; | 10 - Total Intensity; |
| 4 - Total Area (in pixels); | 11 - Avg. Intensity; |
| 5 - Ratio of MajorAxis/MinorAxis; | 12 - Median Intensity; |
| 6 - Eccentricity; | 13 - Intensity of 20% bright pixel |
| 7 - EquivDiameter; | 14 - Intensity of 80% bright pixel |

20 A fragment of an output for a single file, containing 9 images of cells stained for DNA and acquired with a 10x objective. A montage image that was used as a source to generate data in A is presented in Fig. 17.

The same program also summarizes measurements across many files and performs statistical analysis of the summary data. It creates a summary file with
25 the following data:

- | | |
|--|---------------------------------|
| 1 - Image file number; | |
| 2 - Average object Area (in pixels); | 3 - STD (standard deviation) of |
| 2; | |
| 30 4 - Avg. of Ratio of MajorAxis/MinorAxis; | 5 - STD of 4; |
| 6 - Avg. Eccentricity; | 7 - STD of 6; |
| 8 - Avg. EquivDiameter; | 9 - STD of 8; |
| 10 - Avg. of Solidity; | 11 - STD of 10; |

- | | |
|---|----------------|
| 12 - Avg. of Extent; | 13 - STD of 11 |
| 14 - Avg. of objects Total Intensity; | 15 - STD of 14 |
| 16 - Avg. of objects Avg Intensity; | 16 - STD of 15 |
| 18 - Avg. of objects Median intensity; | 19 - STD of 18 |
| 20 - Avg. of objects intensity of 20% bright pixel; | 21 - STD of 19 |
| 22 - Avg. of objects intensity of 80% bright pixel; | 23 - STD of 21 |

An example of summary output obtained by running AnalyseDNA against 10 montage files also is shown in Appendix B.

10 A seventh experiment was conducted in order to use sequence analysis algorithms to analyze features of cell images. In this experiment, HeLa cells were treated for 24 hours with several different compounds, and then fixed, and stained with a fluorescent DNA dye. One image of these cells was acquired for each of the treatments and morphometric parameters and features were measured:

15 Resulting measurements were arranged into a string of numbers and reduced to a pseudo- nucleic acid sequence using following rules: At any given position in the sequence a number was substituted by "t" (a code for thymidine) if its value is among highest 25% of the values at the corresponding position in the data set, "g" if it is between 50% and 25%, "c" if it is between 75% and 50%, and "a" if it
20 belongs to lowest 25% of values. Thus one descriptor or sequence was generated per treatment as illustrated in Fig. 18.

Resulting sequences were clustered using an AlignX module commercial software package Vector NTI (<http://informaxinc.com>), which uses a Neighbor Joining algorithm for sequence clustering.

25 The resulting dendrogram is presented in Fig 18. On the dendrogram the closest "leafs" correspond to the closest pseudo-sequences. Interestingly, compounds with similar mechanisms of action cluster together on the dendrogram. Another example of the generation of pseudo-sequences and clustering is shown in Fig. 19.

30 In some embodiments, techniques according to the present invention can provide tools for the later stages of drug development such as clinical trial design and patient management. The properties of known drugs such as clinical trial and patient response information will be used in a similar fashion as the pre-clinical

information to provide predictions about the properties of novel compounds. Because the human cell is the locus of drug action, a database containing drug-cell interactions can be able to provide predictive information for this aspect of drug development.

Although the above has generally described the present invention

5 according to specific systems, the present invention has a much broader range of applicability. In particular, the present invention is not limited to a particular kind of data about a cell, but can be applied to virtually any cellular data where an understanding about the workings of the cell is desired. Thus, in some embodiments, the techniques of the present invention could provide information about many

10 different types or groups of cells, substances, and genetic processes of all kinds. Of course, one of ordinary skill in the art would recognize other variations, modifications, and alternatives.

APPENDIX A

EV Table 1.doc

Example of the output of AnalyseDNA.m program
(measurements for a single 3 by 3 montage image)

File#	Subimages	object#	X coord	Y coord	Area	Area ratio	Eccentricity	Equi-Diam	Solidity	Extent	Intensity-ty	Avg. Intens.	Median Intens.	20% pta.	80% pta.
1	1	1	12.3897	132.435	145	1.17293	0.322614	12.3875	0.933567	0.730796	1605	31.7586	31	23	37
1	1	2	14.352	416.032	125	1.40394	0.382471	12.6137	0.905789	0.70125	1408	36.910	31	20	45
1	1	3	20.1053	73.9019	177	1.09845	0.43785	15.0131	0.917098	0.691004	1769	28.2925	29	22	31
1	1	4	21.4134	402.714	92	1.32215	0.31004	1.39928	0.914094	0.747837	3690	93.816	87	67	105
1	1	5	23.0938	184.744	86	1.30887	0.465154	11.0598	0.888489	0.671329	4502	46.1938	49	38	54
1	1	6	30.3252	259.334	204	2.3106	0.50209	16.1953	0.97928	0.715378	6380	30.3709	33	24	37
1	1	7	33.6439	147.522	15	1.4584	0.71456	10.6451	0.970933	0.711667	4225	47.4319	50	38	54
1	1	8	35.0411	16.2728	116	1.23174	0.61485	12.8213	0.929526	0.718978	5435	37.089	40	29	44
1	1	9	37.746	346.031	41	1.4042	0.89512	7.7578	0.87037	0.652718	6657	141.853	142	113	171
1	1	10	49.1078	170.004	232	1.50491	0.51127	17.187	0.852911	0.70303	9832	47.3793	45	33	51
1	1	11	51.0749	124.534	221	1.57064	0.69395	16.7746	0.974666	0.665233	7040	31.1552	33	25	37
1	1	12	52.7755	44.2392	147	1.33427	0.60301	12.4409	0.907407	0.706731	4745	32.415	34	24	39
1	1	13	53.1644	286.854	171	2.27275	0.87953	16.7553	0.877418	0.706612	9378	31.4021	34	41	48
1	1	14	58.0229	282.272	206	1.92782	0.83494	16.1953	0.933167	0.64375	7137	31.4156	37	28	41
1	1	15	61.0448	227.174	108	1.73893	0.61809	11.7265	0.915234	0.701789	4644	42	33	51	
1	1	16	64.1714	232.101	313	1.11194	0.437266	20.0267	0.75	0.526756	13151	40.0984	50	36	43
1	1	17	65.1109	402.411	220	1.70147	0.80904	16.7366	0.970502	0.667059	8809	44.3844	46	35	51
1	1	18	72.626	184.854	123	1.73878	0.626593	15.3456	0.91132	0.649123	4824	22.1027	35	28	39
1	1	19	76.7377	209.27	172	1.63279	0.79198	19.7366	0.822501	0.622231	14359	38.3577	41	30	47
1	1	20	77.4868	132.513	304	1.47413	0.794616	12.4634	0.91048	0.739394	4183	47.5784	50	38	53
1	1	21	81.6788	53.5012	117	1.62713	0.626593	15.3456	0.86554	0.62657	4886	40.0513	43	31	47
1	1	22	88.2792	201.534	373	2.17388	0.807916	21.7956	0.843191	0.531339	16109	41.1873	48	34	52
1	1	23	88.1765	341.976	85	1.20799	0.560991	10.4031	0.842489	0.700323	4389	33.9882	37	27	41
1	1	24	90.1408	176.231	143	1.43573	0.717545	11.4935	0.910789	0.794144	4878	34.1119	35	27	41
1	1	25	91.6329	318.824	170	1.34853	0.62669	14.7123	0.919721	0.623323	4389	33.9882	37	27	41
1	1	26	97.7404	211.793	280	1.93233	0.862110	19.3482	0.9	0.606116	10661	37.0143	39	28	43
1	1	27	98.5841	320.363	113	1.09335	0.413609	11.9948	0.894925	0.658638	4340	40.1377	42	32	47
1	1	28	98.8092	248.402	310	1.2774	0.622219	12.2573	0.916755	0.762034	4873	31.2564	43	31	47
1	1	29	103.033	93.2779	222	1.48415	0.738927	12.4634	0.915736	0.613333	4663	48.2213	40	31	47
1	1	30	103.47	155.507	234	1.3708	0.33401	13.0619	0.917608	0.681118	4895	38.7881	41	32	47
1	1	31	105.356	87.1271	110	1.90239	0.850869	12.2573	0.905692	0.694118	4895	38.7881	41	32	47
1	1	32	121.23	285.08	374	1.70319	0.809514	20.3734	0.90533	0.626523	15644	48.4636	50	37	52
1	1	33	125.532	170.445	141	1.50465	0.737013	13.3988	0.927632	0.734235	4139	45.3557	49	37	52
1	1	34	128.88	60.3355	152	1.73689	0.822208	13.9116	0.821212	0.747477	4815	45.2303	47	36	51
1	1	35	137.082	128.083	244	1.845	0.840375	18.6033	0.784607	0.529462	8810	34.8797	37	28	43
1	1	36	130.202	611.5	164	1.19276	0.34507	14.1303	0.927113	0.788142	1337	44.3718	47	36	51
1	1	37	132.429	352.545	187	1.39705	0.682312	15.4204	0.933	0.788145	3277	27.8519	27	27	34
1	1	38	132.429	16.4154	13	1.19251	0.30579	4.01843	0.927571	0.6125	94	7.32017	7	11	11
1	1	39	134.571	209.099	101	1.12013	0.450546	11.3401	0.870781	0.617436	7292	72.2019	71	58	88
1	1	40	134.555	33.0909	33	1.21149	0.364508	6.18204	0.916663	0.782714	1849	28.0406	40	41	75
1	1	41	140.893	102.008	121	1.47187	0.734306	12.4322	0.916663	0.732332	4888	28.7255	40	41	75
1	1	42	141.864	59.8189	272	1.35119	0.821825	18.5087	0.923897	0.732332	10814	31.7574	41	32	48
1	1	43	143.093	438.135	161	1.3071	0.64187	18.2715	0.923897	0.732332	10814	31.7574	41	32	48
1	1	44	143.093	254.324	224	1.08008	0.371873	16.8188	0.923897	0.732332	10814	31.7574	41	32	48
1	1	45	151.46	34.324	141	1.39135	0.652793	13.2888	0.923897	0.732332	10814	31.7574	41	32	48
1	1	46	155.680	178.546	141	1.71551	0.815531	17.814	0.923897	0.732332	10814	31.7574	41	32	48
1	1	47	140.175	342.351	148	1.42126	0.710545	15.8777	0.923897	0.732332	10814	31.7574	41	32	48
1	1	48	140.175	242.351	148	1.42126	0.710545	15.8777	0.923897	0.732332	10814	31.7574	41	32	48
1	1	49	140.175	134.267	217	1.71718	0.920008	15.8777	0.923897	0.732332	10814	31.7574	41	32	48
1	1	50	171.014	356.114	212	1.41189	0.720419	16.8125	0.923897	0.732332	10814	31.7574	41	32	48
1	1	51	175.788	192.983	116	1.13744	0.47768	12.2573	0.923897	0.732332	10814	31.7574	41	32	48
1	1	52	177.181	210.834	121	1.13319	0.487604	12.7182	0.923897	0.732332	10814	31.7574	41	32	48
1	1	53	178.167	40.324	147	1.21277	0.382423	13.4609	0.923897	0.732332	10814	31.7574	41	32	48
1	1	54	182.1	292.176	170	1.37168	0.735981	16.7123	0.923897	0.732332	10814	31.7574	41	32	48
1	1	55	182.1	292.176	170	1.37168	0.735981	16.7123	0.923897	0.732332	10814	31.7574	41	32	48
1	1	56	200.742	93.7418	213	1.81704	0.836936	15.7923	0.918888	0.742424	5030	35.4632	37	20	31
1	1	57	199.130	136.723	91	1.36863	0.683628	16.4882	0.918888	0.742424	5030	35.4632	37	20	31
1	1	58	209.61	183.071	264	1.01653	0.395151	10.7641	0.90039	0.752064	4188	46.032	47	36	55
1	1	59	209.61	183.071	264	1.01653	0.395151	10.7641	0.90039	0.752064	4188	46.032	47	36	55
1	1	60	212.384	346.453	181	1.49829	0.806178	17.1127	0.914502	0.604194	9834	37.2977	39	29	46
1	1	61	212.384	346.453	181	1.49829	0.806178	17.1127	0.914502	0.604194	9834	37.2977	39	29	46
1	1	62	214.560	234.29	183	2.47462	0.923581	15.7165	0.923897	0.732332	10814	31.7574	41	32	48
1	1	63	214.560	234.29	183	2.47462	0.923581	15.7165	0.923897	0.732332	10814	31.7574	41	32	48
1	1	64	217.321	330.321	172	1.72537	0.81091	14.7386	0.918255	0.623084	5078	27.7186	28	21	34
1	1	65	217.321	330.321	172	1.72537	0.81091	14.7386	0.918255	0.623084	5078	27.7186	28	21	34
1	1	66	222.326	157.769	121	1.79272	0.834448	15.9925	0.918255	0.623084	5078	27.7186	28	21	34
1	1	67	249.071	488.948	435	2.20117	0.656743	12.4322	0.916663	0.732332	4832	35.9489	38	29	42
1	1	68	236.144	265.186	167	1.18108	0.738897	14.5819	0.918255	0.623084	5078	27.7186	28	21	34

Page 1 of 16

Page 2 of 16

Page 3 of 16

EV Table 1.doc

1	136	415.89	366.669	299	2.00072	0.331083	19.5113	0.89531	0.48894	1.0053	47.0401	50	38	57
2	137	480.732	106.247	190	1.73179	0.821058	15.3536	0.855056	0.66667	8796	46.2421	50	35	57
3	138	492.316	317.207	195	1.31628	0.679295	15.3536	0.933018	0.71301	8108	27.7332	29	23	57
4	139	10.090	359.078	102	1.21531	0.54819	11.3961	0.918919	0.71301	5508	34	39	59	61
5	140	14.7415	319.365	107	1.8392	0.850132	13.4809	0.800784	0.516671	3099	34.6771	37	25	61
6	141	16.317	274.264	176	1.49674	0.714056	16.9656	0.926116	0.670559	8846	56.0569	40	45	68
7	142	13	274.264	176	1.20918	0.562256	10.2179	0.86172	0.602323	4220	55.8024	56	46	64
8	143	23.1517	333.427	211	1.47212	0.714106	16.2907	0.905579	0.702323	9318	55.2039	47	35	55
9	144	22.7277	333.041	71	1.16531	0.513657	9.70468	0.800551	0.714775	4079	55.1216	56	44	66
10	145	23.7391	366.47	69	1.47156	0.723019	9.7302	0.872116	0.45197	2983	57.7246	58	10	70
11	146	31.3523	50.1478	297	2.54212	0.820487	19.1451	0.838983	0.53102	1184	39.6768	41	31	48
12	147	31.9742	108.832	155	1.60661	0.789975	14.0182	0.937735	0.61964	8887	31.5219	33	24	38
13	148	41.2755	455.739	233	1.9876	0.861616	13.221	0.849025	0.47222	8169	40.2102	42	21	30
14	149	51.2162	275.286	191	2.47167	0.80209	15.3945	0.887116	0.436463	8934	31.1152	51	62	62
15	150	42.4282	262.297	219	1.72956	0.801865	15.4993	0.867205	0.400233	8759	44.3618	46	36	51
16	151	45.7539	171.744	213	2.19727	0.800435	15.4602	0.916667	0.400233	8186	33.1155	46	27	43
17	152	43.4132	488.702	121	1.50032	0.606279	12.4132	0.922664	0.75623	4267	25.2643	31	28	43
18	153	73.3548	80.7162	155	1.30351	0.61193	14.0182	0.922664	0.75623	9210	59.871	41	50	71
19	154	79	279.532	103	1.97089	0.861719	15.7644	0.915	0.714775	4867	21.1181	28	21	33
20	155	80.0786	183.648	210	1.64888	0.781543	16.3518	0.929704	0.714775	7257	31.3571	37	27	41
21	156	75.6586	291.833	141	1.23467	0.602919	13.2988	0.937735	0.713233	4330	31.7533	36	27	40
22	157	80.1048	310.148	141	1.23467	0.602919	13.2988	0.937735	0.713233	4330	31.7533	36	27	40
23	158	86.0366	102.431	161	1.35541	0.674732	13.4651	0.935373	0.713233	4607	37.2937	36	25	39
24	159	90.2161	139.787	171	1.67033	0.800986	14.8843	0.970635	0.651135	4590	26.356	29	21	32
25	160	85.1818	504.591	27	1.58172	0.771787	5.29237	0.916667	0.723233	1517	44.3945	48	64	76
26	161	80.4814	71.4186	43	1.48076	0.646835	13.2988	0.916667	0.723233	1517	44.3945	48	64	76
27	162	95.9043	437.878	188	1.46536	0.731051	15.1716	0.930693	0.705167	4330	31.2319	23	18	28
28	163	102.775	259.464	271	2.75576	0.836369	10.5715	0.73842	0.531371	10362	30.2362	40	30	47
29	164	101.714	331.383	209	1.34189	0.60653	11.7608	0.933229	0.778371	4566	41.8899	44	32	52
30	165	102.831	381.4	105	1.48432	0.801731	11.4641	0.933043	0.673071	4538	43.219	49	34	53
31	166	102.731	26.2077	182	1.16701	0.735179	15.2273	0.938144	0.722222	5125	29.2132	31	24	35
32	167	104.742	241.391	232	1.53467	0.735179	15.2273	0.938144	0.722222	5125	29.2132	31	24	35
33	168	101.802	237.816	120	1.2205	0.58271	12.1655	0.90571	0.722222	4831	37.1615	37	28	46
34	169	113.68	339.312	128	1.2113	0.584315	12.7682	0.941176	0.715105	4903	35.1717	37	28	46
35	170	117.636	420.502	145	1.53776	0.759484	13.4923	0.91395	0.710784	3953	27.131	28	21	24
36	171	119.99	274.808	104	1.58238	0.775046	11.5073	0.91395	0.710784	3953	27.131	28	21	24
37	172	123.631	445.233	160	2.13328	0.832025	14.273	0.85515	0.446463	3953	27.131	28	21	24
38	173	125.541	492.442	123	1.47198	0.681226	13.0131	0.9001	0.720769	3821	20.7263	35	19	31
39	174	129.158	285.54	114	1.81038	0.823599	12.0131	0.91395	0.710784	3953	27.131	28	21	24
40	175	135.442	301.207	80	1.7702	0.823554	10.0925	0.888716	0.711716	4310	52.9235	55	42	63
41	176	144.798	321.454	216	1.78966	0.823272	16.5937	0.925065	0.75	8028	37.1663	34	20	44
42	177	146.098	351.596	114	1.78966	0.823272	16.5937	0.925065	0.75	8028	37.1663	34	20	44
43	178	147.37	232.151	119	1.2321	0.571819	12.3092	0.922181	0.712323	8726	49.1281	72	52	83
44	179	150.137	435.098	102	1.36059	0.678555	11.3961	0.910711	0.71301	3854	31.7843	40	31	45
45	180	154.435	412.167	104	1.47096	0.733370	11.7255	0.915354	0.701996	4156	34.4815	40	30	46
46	181	158.445	295.732	254	2.29556	0.90072	13.8934	0.84104	0.529167	9778	34.4961	40	31	48
47	182	159.39	306.357	201	1.61833	0.793481	15.9975	0.918884	0.705353	5892	34.7861	36	27	42
48	183	165.739	277.091	141	1.72932	0.820519	13.1106	0.912077	0.723463	6195	34.5532	43	32	48
49	184	166.137	300.408	161	1.65349	0.784846	13.1319	0.918884	0.705353	5892	34.5532	43	32	48
50	185	170.119	432.128	119	1.62922	0.711427	12.3092	0.915353	0.701996	10321	34.1045	41	28	42
51	186	174.646	471.702	131	1.28262	0.678846	13.2988	0.915353	0.701996	10321	34.1045	41	28	42
52	187	175.474	172.404	134	1.23535	0.678846	13.2988	0.915353	0.701996	10321	34.1045	41	28	42
53	188	182.262	71.0719	132	1.61934	0.709554	13.2988	0.921887	0.75	4933	22.2723	34	26	38
54	189	186.416	428.314	114	2.16718	0.848278	13.2988	0.89161	0.531429	4506	31.2937	33	23	35
55	190	196.816	172.408	232	1.59008	0.678846	13.2988	0.901321	0.636957	21249	72.3222	71	51	89
56	191	191.99	210.359	286	1.78224	0.827754	13.0918	0.925366	0.723233	10274	35.2201	37	29	41
57	192	202.502	293.631	232	2.40146	0.907229	13.2117	0.91398	0.631466	10274	35.2201	37	29	41
58	193	206.034	10.4413	232	1.32324	0.634852	15.4115	0.913709	0.739197	9631	43.4081	46	34	53
59	194	218.937	319.989	269	1.78272	0.827856	13.5088	0.916667	0.67398	10198	37.3108	40	30	45
60	195	217.314	109.796	250	1.92424	0.843337	17.8412	0.916667	0.548246	9609	38.1316	39	29	48
61	196	215.452	102.911	157	1.45463	0.727025	14.1286	0.918993	0.548246	9609	38.1316	39	29	48
62	197	221.324	468.737	304	1.61216	0.831019	19.871	0.91106	0.603175	20032	65.9603	70	51	81
63	198	224.342	36.0972	169	1.27551	0.615353	13.7736	0.91325	0.776002	4908	30.8638	32	25	36
64	199	228.818	451.459	159	1.38781	0.623392	17.2712	0.921619	0.719157	4908	30.8638	32	25	36
65	200	236.586	365.853	237	1.34178	0.646762	17.2712	0.921619	0.719157	4908	30.8638	32	25	36
66	201	235.133	12.3772	97	1.30129	0.736418	11.1132	0.941748	0.719157	4908	30.8638	32	25	36
67	202	238.177	54.3398	181	1.30129	0.736418	11.1132	0.941748	0.719157	4908	30.8638	32	25	36
68	203	242.074	191.748	189	1.63183	0.811115	15.1808	0.928205	0.719157	4908	30.8638	32	25	36
69	204	245.428	301.019	225	1.72034	0.816331	12.7162	0.911286	0.719157	4908	30.8638	32	25	36
70	205	246.148	447.484	155	1.7135	0.812141	13.0182	0.910394	0.621174	8887	31.2129	32	23	37
71	206	247.201	34.1051	110	1.47373	0.634906	14.0182	0.921619	0.719157	4908	30.8638	32	25	36
72	207	248.393	42.1091	129	1.53133	0.636035	13.2018	0.921619	0.719157	4908	30.8638	32	25	36
73	208	249.481	285.887	71	1.38461	0.631716	8.77205	0.881353	0.651664	8191	109.213	108	84	132
74	209	249.481	145.982	59	1.46172	0.601	6.46774	0.881353	0.651664	8191	109.213	108	84	132
75	210	249.481	145.982	59	1.46172	0.601	6.46774	0.881353	0.651664	8191	109.213	108	84	132
76	211	249.481	145.982	59	1.46172	0.601	6.46774	0.881353	0.651664	8191	109.213	108	84	132
77	212	249.481	145.982	59	1.46172	0.601	6.46774	0.881353	0.651664	8191	109.213	108	84	132
78	213	249.481	145.982	59	1.46172	0.601	6.46774	0.881353	0.651664	8191	109.213	108	84	132
79	214	249.481	145.982	59	1.46172	0.601	6.46774	0.881353	0.6					

1	1	270.921	433.417	177	1.27866	0.674112	13.0121	0.217098	0.713537	9719	53	42	59
2	2	271.454	412.443	221	1.27898	0.615564	14.7146	0.1841	0.613649	9712	52	41	58
3	3	272.022	395.475	265	1.28014	0.561774	16.2237	0.160737	0.561026	9705	51	40	57
4	4	272.621	381.164	308	1.28161	0.512535	17.5287	0.141611	0.510025	9698	50	39	56
5	5	273.243	368.635	346	1.28335	0.468295	18.6858	0.125068	0.466467	9691	49	38	55
6	6	273.888	357.604	383	1.28535	0.428094	19.7317	0.110723	0.426109	9684	48	37	54
7	7	274.557	347.661	419	1.28762	0.391817	20.6817	0.100001	0.390151	9677	47	36	53
8	8	275.250	338.604	454	1.29017	0.358017	21.5516	0.090883	0.356422	9670	46	35	52
9	9	275.966	330.331	489	1.29299	0.327226	22.3517	0.083083	0.325722	9663	45	34	51
10	10	276.704	322.741	524	1.29607	0.298972	23.0818	0.076425	0.297972	9656	44	33	50
11	11	277.464	315.741	559	1.30041	0.272726	23.7519	0.070825	0.270025	9649	43	32	49
12	12	278.246	309.241	594	1.30501	0.248972	24.3719	0.066125	0.244972	9642	42	31	48
13	13	279.050	303.141	629	1.31087	0.227226	24.9519	0.062325	0.222325	9635	41	30	47
14	14	279.876	297.341	664	1.31700	0.207226	25.4919	0.059325	0.209325	9628	40	29	46
15	15	280.724	291.741	699	1.32340	0.188672	26.0019	0.057025	0.189025	9621	39	28	45
16	16	281.594	286.241	734	1.33007	0.171226	26.4819	0.055325	0.171226	9614	38	27	44
17	17	282.486	280.841	769	1.33699	0.154872	26.9319	0.054125	0.154872	9607	37	26	43
18	18	283.400	275.541	804	1.34416	0.139572	27.3519	0.053325	0.139572	9600	36	25	42
19	19	284.336	270.341	839	1.35157	0.125272	27.7419	0.052925	0.125272	9593	35	24	41
20	20	285.294	265.141	874	1.35922	0.111872	28.1019	0.052825	0.111872	9586	34	23	40
21	21	286.274	260.041	909	1.36710	0.099272	28.4319	0.053025	0.099272	9579	33	22	39
22	22	287.276	255.041	944	1.37521	0.087272	28.7319	0.053525	0.087272	9572	32	21	38
23	23	288.300	250.041	979	1.38354	0.075872	29.0019	0.054325	0.075872	9565	31	20	37
24	24	289.346	245.041	1014	1.39209	0.065072	29.2419	0.055325	0.065072	9558	30	19	36
25	25	290.414	240.041	1049	1.40086	0.054872	29.4519	0.056525	0.054872	9551	29	18	35
26	26	291.504	235.041	1084	1.40984	0.045272	29.6319	0.057925	0.045272	9544	28	17	34
27	27	292.616	230.041	1119	1.41903	0.036272	29.7819	0.059525	0.036272	9537	27	16	33
28	28	293.750	225.041	1154	1.42842	0.027872	29.9019	0.061325	0.027872	9530	26	15	32
29	29	294.906	220.041	1189	1.43801	0.019972	30.0019	0.063325	0.019972	9523	25	14	31
30	30	296.084	215.041	1224	1.44779	0.012572	30.0819	0.065525	0.012572	9516	24	13	30
31	31	297.284	210.041	1259	1.45776	0.005672	30.1419	0.067925	0.005672	9509	23	12	29
32	32	298.506	205.041	1294	1.46791	0.000272	30.1819	0.070525	0.000272	9502	22	11	28
33	33	299.750	200.041	1329	1.47824	0.000000	30.2019	0.073325	0.000000	9495	21	10	27
34	34	301.016	195.041	1364	1.48874	0.000000	30.2119	0.076325	0.000000	9488	20	9	26
35	35	302.304	190.041	1399	1.49941	0.000000	30.2119	0.079525	0.000000	9481	19	8	25
36	36	303.614	185.041	1434	1.51024	0.000000	30.1919	0.082925	0.000000	9474	18	7	24
37	37	304.946	180.041	1469	1.52122	0.000000	30.1519	0.086525	0.000000	9467	17	6	23
38	38	306.300	175.041	1504	1.53235	0.000000	30.0919	0.090325	0.000000	9460	16	5	22
39	39	307.676	170.041	1539	1.54362	0.000000	30.0119	0.094325	0.000000	9453	15	4	21
40	40	309.074	165.041	1574	1.55503	0.000000	29.9119	0.098525	0.000000	9446	14	3	20
41	41	310.494	160.041	1609	1.56658	0.000000	29.7919	0.102925	0.000000	9439	13	2	19
42	42	311.936	155.041	1644	1.57827	0.000000	29.6519	0.107525	0.000000	9432	12	1	18
43	43	313.400	150.041	1679	1.59009	0.000000	29.4919	0.112325	0.000000	9425	11	0	17
44	44	314.886	145.041	1714	1.60204	0.000000	29.3119	0.117325	0.000000	9418	10	0	16
45	45	316.394	140.041	1749	1.61412	0.000000	29.1119	0.122525	0.000000	9411	9	0	15
46	46	317.924	135.041	1784	1.62632	0.000000	28.8919	0.127925	0.000000	9404	8	0	14
47	47	319.476	130.041	1819	1.63864	0.000000	28.6519	0.133525	0.000000	9397	7	0	13
48	48	321.050	125.041	1854	1.65108	0.000000	28.3919	0.139325	0.000000	9390	6	0	12
49	49	322.646	120.041	1889	1.66364	0.000000	28.1119	0.145325	0.000000	9383	5	0	11
50	50	324.264	115.041	1924	1.67632	0.000000	27.8119	0.151525	0.000000	9376	4	0	10
51	51	325.904	110.041	1959	1.68912	0.000000	27.4919	0.157925	0.000000	9369	3	0	9
52	52	327.566	105.041	1994	1.70204	0.000000	27.1519	0.164525	0.000000	9362	2	0	8
53	53	329.250	100.041	2029	1.71508	0.000000	26.7919	0.171325	0.000000	9355	1	0	7
54	54	330.956	95.041	2064	1.72824	0.000000	26.4119	0.178325	0.000000	9348	0	0	6
55	55	332.684	90.041	2099	1.74152	0.000000	26.0119	0.185525	0.000000	9341	0	0	5
56	56	334.434	85.041	2134	1.75492	0.000000	25.5919	0.192925	0.000000	9334	0	0	4
57	57	336.206	80.041	2169	1.76844	0.000000	25.1519	0.200525	0.000000	9327	0	0	3
58	58	337.999	75.041	2204	1.78208	0.000000	24.6919	0.208325	0.000000	9320	0	0	2
59	59	339.814	70.041	2239	1.79584	0.000000	24.2119	0.216325	0.000000	9313	0	0	1
60	60	341.650	65.041	2274	1.80972	0.000000	23.7119	0.224525	0.000000	9306	0	0	0
61	61	343.506	60.041	2309	1.82372	0.000000	23.1919	0.232925	0.000000	9299	0	0	0
62	62	345.382	55.041	2344	1.83784	0.000000	22.6519	0.241525	0.000000	9292	0	0	0
63	63	347.278	50.041	2379	1.85208	0.000000	22.0919	0.250325	0.000000	9285	0	0	0
64	64	349.194	45.041	2414	1.86644	0.000000	21.5119	0.259325	0.000000	9278	0	0	0
65	65	351.130	40.041	2449	1.88092	0.000000	20.9119	0.268525	0.000000	9271	0	0	0
66	66	353.086	35.041	2484	1.89552	0.000000	20.2919	0.277925	0.000000	9264	0	0	0
67	67	355.062	30.041	2519	1.91024	0.000000	19.6519	0.287525	0.000000	9257	0	0	0
68	68	357.058	25.041	2554	1.92508	0.000000	18.9919	0.297325	0.000000	9250	0	0	0
69	69	359.074	20.041	2589	1.93992	0.000000	18.3119	0.307325	0.000000	9243	0	0	0
70	70	361.110	15.041	2624	1.95484	0.000000	17.6119	0.317525	0.000000	9236	0	0	0
71	71	363.166	10.041	2659	1.96984	0.000000	16.8919	0.327925	0.000000	9229	0	0	0
72	72	365.242	5.041	2694	1.98492	0.000000	16.1519	0.338525	0.000000	9222	0	0	0
73	73	367.338	0.041	2729	1.99992	0.000000	15.3919	0.349325	0.000000	9215	0	0	0
74	74	369.454	0.041	2764	2.01492	0.000000	14.6119	0.360325	0.000000	9208	0	0	0
75	75	371.590	0.041	2799	2.02992	0.000000	13.8119	0.371525	0.000000	9201	0	0	0
76	76	373.746	0.041	2834	2.04492	0.000000	13.0019	0.382925	0.000000	9194	0	0	0
77	77	375.922	0.041	2869	2.05992	0.000000	12.1719	0.394525	0.000000	9187	0	0	0
78	78	378.118	0.041	2904	2.07492	0.000000	11.3219	0.406325	0.000000	9180	0	0	0
79	79	380.334	0.041	2939	2.08992	0.000000	10.4519	0.418325	0.000000	9173	0	0	0
80	80	382.570	0.041	2974	2.10492	0.000000	9.5619	0.430525	0.000000	9166	0	0	0
81	81	384.826	0.041	3009	2.11992	0.000000	8.6519	0.442925	0.000000	9159	0	0	0
82	82	387.102	0.041	3044	2.13492	0.000000	7.7219	0.455525	0.000000	9152	0	0	0
83	83	389.398	0.041	3079	2.14992	0.000000	6.7719	0.468325	0.000000	9145	0	0	0
84	84	391.714	0.041	3114	2.16492	0.000000	5.8019	0.481325	0.000000	9138	0	0	0
85	85	394.050	0.041	3149	2.17992	0.000000	4.8119	0.494525	0.000000	9131	0	0	0
86	86	396.406	0.041	3184	2.19492	0.000000	3.8019	0.507925	0.000000	9124	0	0	0

Page 7 of 16

[illegible]

Page 8 of 16

[illegible]

EV Table 1.doc

1	1	132.384	170.016	250	1.31824	0.813119	17.8412	0.329168	0.457195	3096	20.391	71	23
2	1	138.47	178.326	115	1.31792	0.812329	17.1005	0.331671	0.451497	1461	36.1826	78	41
3	1	135.391	231.210	135	1.30495	0.810518	17.5573	0.310714	0.480133	4600	78.7362	28	35
4	1	129.828	361.001	99	1.31318	0.750701	17.2212	0.306581	0.466523	9241	81.1111	45	51
5	1	135.662	376.305	196	1.40408	0.804616	15.7973	0.309571	0.476571	3972	91.104	10	39
6	1	136.281	415.42	171	1.40358	0.78115	14.7355	0.314723	0.470453	1921	21.2811	24	28
7	1	133.941	432	118	1.46552	0.816151	12.2373	0.314723	0.470453	1921	34.1017	36	42
8	1	163.395	40.609	205	2.45432	0.815225	14.3559	0.323623	0.471623	1596	22.5195	74	76
9	1	136.273	497.208	107	2.41308	0.923253	11.672	0.325216	0.423612	4016	22.5195	74	76
10	1	161.448	246.271	181	3.20118	0.609916	13.1008	0.318162	0.709801	8138	44.7323	43	51
11	1	153.392	30.4022	181	1.62704	0.791716	11.4115	0.310516	0.710516	8138	44.7323	43	51
12	1	153.004	138.448	173	1.62704	0.791716	11.4115	0.310516	0.710516	8138	44.7323	43	51
13	1	151.167	311.515	159	1.81167	0.835356	14.2283	0.313793	0.711778	4453	44.7323	43	51
14	1	151.466	42.3155	97	1.81167	0.835356	14.2283	0.313793	0.711778	4453	44.7323	43	51
15	1	163.426	191.575	73	1.35143	0.612771	11.5424	0.305172	0.731666	4936	31.5899	32	40
16	1	161.439	431.371	105	1.30519	0.817937	9.6088	0.305172	0.731666	4936	31.5899	32	40
17	1	161.203	71.1556	91	2.09666	0.810516	10.1554	0.31	0.731666	3980	49.1358	50	40
18	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
19	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
20	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
21	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
22	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
23	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
24	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
25	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
26	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
27	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
28	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
29	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
30	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
31	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
32	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
33	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
34	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
35	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
36	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
37	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
38	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
39	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
40	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
41	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
42	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
43	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
44	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
45	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
46	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
47	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
48	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
49	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
50	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
51	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
52	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
53	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
54	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
55	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
56	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
57	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
58	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
59	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
60	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
61	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
62	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
63	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
64	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
65	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
66	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
67	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
68	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
69	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
70	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
71	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
72	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
73	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
74	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
75	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
76	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
77	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
78	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
79	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
80	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
81	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
82	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
83	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
84	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
85	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	3980	49.1358	50	40
86	1	171.794	405.46	126	1.15273	0.738516	10.8117	0.31	0.731666	39			

EV Table 1.duc

1	106	358.018	312.36	114	1.71044	0.871112	12.0418	0.897628	0.426374	4181	36.4751	38	78	45
2	107	361.714	169.021	129	1.40135	0.781018	12.8159	0.80851	0.461538	7531	58.1166	40	41	71
3	108	370.42	124.784	112	1.12716	0.46271	11.7116	0.493126	0.717493	4218	37.4307	40	20	44
4	109	379.864	101.258	97	1.21577	0.588227	11.1122	0.493126	0.717493	4114	42.7216	44	31	51
5	110	382.727	404.484	107	1.71038	0.810259	15.7061	0.812155	0.811111	8110	44.1497	44	35	51
6	111	388.96	193.188	203	1.81034	0.835333	15.4416	0.812155	0.811111	8110	38.8157	39	29	43
7	112	397.155	337.22	171	1.72584	0.811993	11.4643	0.811993	0.775	8151	50.3116	31	27	41
8	113	398.864	241.461	134	1.42811	0.801733	11.359	0.811993	0.775	8151	22.6164	35	27	41
9	114	401.428	202.432	118	1.51997	0.732098	22.3591	0.846108	0.812155	1864	41.3159	30	22	41
10	115	395.517	232.561	118	2.08819	0.871645	12.7735	0.846108	0.812155	5411	36.3154	38	22	41
11	116	414.693	272.039	384	1.50902	0.748061	22.1114	0.811993	0.775	8151	38.8157	31	22	33
12	117	424.431	70.7462	159	1.63116	0.701051	14.7281	0.811993	0.775	8151	38.8157	31	22	33
13	118	429	16	109	1.21224	0.591449	11.7864	0.811993	0.775	8151	38.8157	31	22	33
14	119	434.508	178.295	132	1.21224	0.591449	11.7864	0.811993	0.775	8151	38.8157	31	22	33
15	120	442.816	371.2	335	1.41131	0.81648	21.4211	0.811993	0.775	8151	38.8157	31	22	33
16	121	435.418	104.209	81	1.47374	0.713559	10.7461	0.811993	0.775	8151	38.8157	31	22	33
17	122	439.948	245.401	111	1.71057	0.811993	11.9119	0.811993	0.775	8151	38.8157	31	22	33
18	123	441.315	472.301	110	1.32152	0.735019	11.0115	0.811993	0.775	8151	38.8157	31	22	33
19	124	447.904	87.8919	114	1.58128	0.735019	11.0115	0.811993	0.775	8151	38.8157	31	22	33
20	125	450.265	484.34	113	1.271601	0.668112	11.9918	0.811993	0.775	8151	38.8157	31	22	33
21	126	454.785	100.813	107	1.271601	0.668112	11.9918	0.811993	0.775	8151	38.8157	31	22	33
22	127	455.11	159.5	100	1.271601	0.668112	11.9918	0.811993	0.775	8151	38.8157	31	22	33
23	128	461.271	117.474	123	1.7029	0.809391	11.0115	0.811993	0.775	8151	38.8157	31	22	33
24	129	472.26	224.255	101	1.271601	0.668112	11.9918	0.811993	0.775	8151	38.8157	31	22	33
25	130	472.405	134.154	114	1.271601	0.668112	11.9918	0.811993	0.775	8151	38.8157	31	22	33
26	131	473.51	347.364	114	1.271601	0.668112	11.9918	0.811993	0.775	8151	38.8157	31	22	33
27	132	479.14	22.1551	108	1.42015	0.710399	11.7104	0.811993	0.775	8151	38.8157	31	22	33
28	133	483.404	40.7073	164	2.01452	0.816305	14.4503	0.811993	0.775	8151	38.8157	31	22	33
29	134	480.291	141.805	151	1.41152	0.707316	14.8843	0.811993	0.775	8151	38.8157	31	22	33
30	135	490.391	331.707	174	1.55716	0.746379	14.8843	0.811993	0.775	8151	38.8157	31	22	33
31	136	497.804	512.497	142	2.40041	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
32	137	496.029	311.5	200	2.05046	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
33	138	496.219	7.97088	95	1.70584	0.871029	11.0115	0.811993	0.775	8151	38.8157	31	22	33
34	139	496.817	238.492	115	1.10135	0.553944	10.4817	0.811993	0.775	8151	38.8157	31	22	33
35	140	501.108	501.455	99	1.62754	0.712166	11.2212	0.811993	0.775	8151	38.8157	31	22	33
36	141	12.1717	59.1515	99	1.62754	0.712166	11.2212	0.811993	0.775	8151	38.8157	31	22	33
37	142	22.2837	46.7188	98	1.32024	0.655562	11.2704	0.811993	0.775	8151	38.8157	31	22	33
38	143	24.3438	315.156	84	1.34015	0.655562	11.2704	0.811993	0.775	8151	38.8157	31	22	33
39	144	31.2359	21.5351	211	1.67385	0.801881	11.5755	0.811993	0.775	8151	38.8157	31	22	33
40	145	31.27	130.32	100	1.20945	0.645736	11.7838	0.811993	0.775	8151	38.8157	31	22	33
41	146	30.458	144.426	131	1.42932	0.715017	12.3153	0.811993	0.775	8151	38.8157	31	22	33
42	147	31.7195	193.978	116	1.42932	0.715017	12.3153	0.811993	0.775	8151	38.8157	31	22	33
43	148	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
44	149	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
45	150	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
46	151	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
47	152	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
48	153	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
49	154	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
50	155	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
51	156	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
52	157	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
53	158	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
54	159	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
55	160	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
56	161	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
57	162	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
58	163	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
59	164	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
60	165	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
61	166	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
62	167	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
63	168	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
64	169	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
65	170	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
66	171	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
67	172	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
68	173	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
69	174	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
70	175	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
71	176	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
72	177	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
73	178	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
74	179	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
75	180	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
76	181	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
77	182	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
78	183	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
79	184	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157	31	22	33
80	185	35.517	379.207	113	1.71056	0.811993	11.7104	0.811993	0.775	8151	38.8157			

Page 12 of 16

EV Table 1.doc

1	1	13.1154	387.567	104	1.42546	0.71237	11.5073	0.90316	0.45325	2355	29.0288	29	31	45
2	2	9.0003	429.712	66	1.2049	0.55191	5.167	0.46	0.72123	3559	66.2485	40	53	
3	3	11.0364	435.810	77	1.17652	0.52646	9.0049	0.52711	0.7	2970	38.9312	40	54	
4	4	14.4166	456.366	70	1.03328	0.35764	9.4019	0.61606	0.7	2970	56.7143	58	66	
5	5	10.3917	475.1216	116	1.763	0.10012	12.0018	0.912	0.72019	3892	34.1404	35	27	
6	6	21.1	472.48	100	1.52812	0.73151	11.2818	0.84456	0.71286	2460	39.4	39	41	
7	7	28.1143	499.143	140	1.52652	0.73556	13.5312	0.90326	0.67077	4072	28.0657	30	22	
8	8	26.0148	432.905	189	1.55415	0.95263	15.5116	0.83628	0.63	12608	66.709	59	51	84
9	9	26.6947	450.616	114	1.70311	0.61665	12.0018	0.91935	0.70749	4004	35.1228	37	28	
10	10	27.688	417.544	225	1.44333	0.721019	12.6157	0.90577	0.69141	7138	57.104	60	41	
11	11	27.9073	348.684	133	1.21356	0.56442	13.0131	0.81724	0.73689	4708	31.6391	32	25	
12	12	28.6993	66.407	113	1.21116	0.40157	0.31803	0.93226	0.721359	4075	36.0619	37	29	
13	13	30.1657	380.429	196	1.40157	0.31803	15.7973	0.922591	0.745714	4063	28.352	30	22	
14	14	36.7162	311.455	143	1.07169	0.271697	13.4935	0.922591	0.745714	4063	28.352	30	22	
15	15	48.0119	37.619	252	1.38006	0.69066	17.8125	0.62623	0.78727	11627	63.2452	45	36	
16	16	65.7422	123.423	122	1.20508	0.42506	12.4638	0.917263	0.782051	4462	36.2131	39	30	
17	17	64.7636	203.731	134	1.40436	0.761967	13.4619	0.937063	0.60473	4152	30.9651	32	25	
18	18	50.7117	48.5565	212	1.49816	0.761967	13.4619	0.917769	0.62181	8604	41.3283	42	32	
19	19	53.3043	661.719	92	1.47116	0.72655	10.823	0.910651	0.69571	3441	39.5741	39	31	
20	20	64.0732	162.192	478	1.03576	0.266543	24.67	0.910651	0.733385	14880	31.1271	32	24	
21	21	64.4333	121.608	120	1.71461	0.557792	12.3608	0.95372	0.710059	4070	33.9147	35	27	
22	22	68.2803	170.141	223	1.5985	0.780155	14.7146	0.634453	0.735617	7881	35.7059	37	28	
23	23	64.7156	219.538	156	1.44808	0.763371	14.0935	0.923072	0.737819	4388	27.5513	28	21	
24	24	66.2917	50.0766	144	1.23123	0.439376	12.5408	0.917349	0.6	9206	36.1528	27	28	
25	25	67.2043	477	93	1.31235	0.460394	10.8813	0.90913	0.65035	3728	40.2043	41	26	
26	26	72.4838	297.915	117	1.21604	0.481693	15.5126	0.917416	0.61937	3343	29.3175	30	23	
27	27	72.382	601.974	189	2.01463	0.806551	12.0018	0.911565	0.67917	4152	30.9851	33	24	
28	28	76.4316	134.731	134	1.20384	0.406551	12.0018	0.911565	0.67917	4152	30.9851	33	24	
29	29	76.6636	108.148	107	1.23861	0.590113	11.672	0.90576	0.746252	4122	38.5234	39	30	
30	30	85.2585	160.102	147	2.77423	0.706143	13.6808	0.90164	0.66482	4231	28.7823	29	23	
31	31	91.282	74.7111	123	1.16581	0.317778	11.2948	0.90164	0.724255	4168	36.485	38	28	
32	32	91.6089	208.4	103	1.45729	0.742048	22.4521	0.944339	0.64953	14872	26.3022	37	28	
33	33	90.5396	160.101	109	1.37151	0.742048	11.7806	0.934163	0.64881	4033	27.1923	38	30	
34	34	96.3458	647.375	273	1.21869	0.425109	18.4719	0.937816	0.386464	12658	51.1172	32	25	
35	35	99.1137	467.323	299	1.62462	0.761178	19.3115	0.931204	0.684211	5700	32.0067	32	25	
36	36	96.2714	37.536	277	2.86001	0.601255	18.78	0.821168	0.50272	7730	27.1061	29	21	
37	37	96.7034	166.452	93	1.26096	0.401152	10.6117	0.865714	0.50258	3856	41.4621	42	32	
38	38	100.312	493.449	203	2.12662	0.882736	14.1356	0.931111	0.719064	4428	22.5756	23	17	
39	39	112.24	441.36	150	1.6122	0.764417	13.6198	0.931677	0.751576	7793	33.9323	53	40	
40	40	111.812	222.761	117	1.59263	0.45094	12.2053	0.906677	0.764708	4458	38.1026	39	31	
41	41	113.866	346.431	149	1.39804	0.698828	13.7728	0.91411	0.761036	4233	28.4096	30	21	
42	42	120.83	411.358	106	1.21794	0.45137	11.6176	0.923823	0.615203	3933	37.0064	37	29	
43	43	129.229	192.142	64	1.15603	0.30172	10.3418	0.90326	0.694215	2693	32.0395	32	25	
44	44	137.15	456.365	232	1.89597	0.649596	17.9128	0.954237	0.672	9367	27.23	38	28	
45	45	134.843	126.374	203	2.06223	0.774693	18.7013	0.914695	0.600804	8260	27.1475	26	21	
46	46	141.141	215.14	219	1.50612	0.746532	20.1392	0.937749	0.693476	11477	35.9781	36	28	
47	47	137.059	242.738	119	1.32113	0.613111	12.3092	0.915385	0.777727	3564	20.1176	31	24	
48	48	145.167	17.8986	207	1.73569	0.62139	16.2345	0.921107	0.704082	6382	41.4549	44	32	
49	49	153.278	417.183	113	1.65216	0.796105	12.1005	0.927119	0.60473	4040	35.1308	37	26	
50	50	159.463	276.456	136	2.03914	0.741056	13.1559	0.937931	0.741253	4230	31.1029	37	25	
51	51	159.706	91.2327	245	2.00914	0.617377	11.4619	0.926528	0.733846	6913	36.3796	37	28	
52	52	159.59	393.024	83	1.4084	0.70128	10.28	0.912088	0.691667	3838	46.241	47	37	
53	53	166.993	500.36	175	1.63183	0.806417	14.2731	0.91623	0.641362	4460	25.6	26	20	
54	54	170.932	257.131	375	1.34635	0.665375	21.851	0.910194	0.708465	14547	38.762	39	31	
55	55	177.162	34.2375	167	1.32019	0.753191	14.5819	0.91566	0.701611	5095	30.309	32	24	
56	56	178.6	387.358	65	1.20726	0.560236	10.9811	0.896204	0.716497	3426	40.2121	40	32	
57	57	182.352	625.777	224	2.70027	0.928932	16.818	0.951218	0.515119	6097	40.2088	40	31	
58	58	179.371	458.119	101	1.32641	0.578918	13.3101	0.916506	0.751332	4011	39.7128	41	33	
59	59	183.359	160.741	129	1.22747	0.389013	12.7462	0.946148	0.620323	2772	21.4553	21	17	
60	60	183.464	505.145	110	1.46155	0.844013	11.4315	0.807057	0.603384	4158	37.7816	39	27	
61	61	197.016	126.252	223	1.43265	0.790527	16.8503	0.90928	0.726028	4777	26.2387	39	30	
62	62	191.661	214.072	138	2.24327	0.891257	12.2555	0.907495	0.746667	4036	26.2464	30	22	
63	63	199.641	455.08	113	1.71204	0.811682	11.9918	0.91129	0.814848	4036	35.7168	36	27	
64	64	207.119	262.204	226	1.93002	0.653702	14.9623	0.896643	0.724239	7354	34.3097	36	27	
65	65	203.414	35.0171	117	1.26158	0.609649	12.2053	0.92126	0.75571	3935	32.9744	35	26	
66	66	206.23	258.306	178	1.36462	0.680437	15.0545	0.92228	0.706349	6679	37.5125	39	30	
67	67	210.404	15.8125	128	1.45902	0.781174	12.7642	0.901404	0.711111	4017	31.3826	32	24	
68	68	220.184	93.4399	161	1.13172	0.472599	24.2273	0.931313	0.770903	16442	60.0073	42	30	
69	69	225.176	214.767	163	1.11692	0.46118	14.4062	0.926136	0.724641	4374	26.8346	28	21	
70	70	230.978	235.801	116	1.31912	0.65216	13.4343	0.918239	0.695236	4334	28.2131	29	22	
71	71	236.473	495.272	316	2.30331	0.809331	20.9891	0.800826	0.619711	10802	21.2197	33	23	
72	72	235.102	132.659	255	1.71464	0.820788	16.0188	0.897687	0.613336	4318	34.9725	36	27	
73	73	231.124	391.719	89	1.45186	0.724578	10.4551	0.89689	0.706464	4214	46.4719	47	36	
74	74	242.141	156.406	251	2.26019	0.896797	17.8766	0.820261	0.602368	4132	36.3625	38	28	
75	75	239.34	189.286	162	1.7197	0.820568	14.3619	0.95011	0.692301	4164	25.8372	26	20	
76	76	239.314	21.4651	138	1.71073	0.625268	13.2555	0.92	0.651143	4080	28.5552	30	23	

Page 13 of 16

EV Table 1.doc

1	21.2188	364.123	64	1.29341	0.634372	0.92703	0.476712	0.127273	3642	56.9063	57	66
2	24.0233	633.466	129	1.82865	0.873231	12.8159	0.621429	0.661536	3872	30.0135	30	23
3	35.6126	114.393	140	1.45761	0.72155	13.3512	0.909091	0.729167	4313	30.8071	31	23
4	39.8172	251.771	206	1.6419	0.74483	14.1933	0.919403	0.477632	9014	43.7573	45	35
5	38.3562	226.443	178	1.66005	0.321795	15.8195	0.927003	0.791111	9196	51.4529	52	38
6	39.3667	41.32	150	1.17673	0.327089	17.8198	0.920165	0.745306	1709	31.3233	32	23
7	36.2345	349.388	85	1.17688	0.472453	16.4031	0.894737	0.702339	1960	46.4235	49	39
8	45.223	24.1727	139	1.40448	0.703172	13.3034	0.934631	0.444274	4279	30.7842	32	25
9	57.3314	101.148	187	1.77115	0.624976	15.4304	0.911193	0.487857	4410	22.5018	24	18
10	55.8904	42.5274	146	1.41519	0.748473	12.4513	0.910051	0.453238	4288	79.4316	30	22
11	65.4198	176.378	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
12	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
13	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
14	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
15	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
16	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
17	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
18	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
19	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
20	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
21	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
22	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
23	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
24	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
25	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
26	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
27	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
28	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
29	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
30	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
31	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
32	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
33	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
34	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
35	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
36	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
37	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
38	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
39	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
40	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
41	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
42	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
43	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
44	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
45	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
46	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
47	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
48	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
49	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
50	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
51	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
52	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
53	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
54	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
55	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
56	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
57	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
58	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
59	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
60	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
61	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23
62	61.4418	176.366	126	1.31497	0.620718	12.466	0.9	0.453208	4098	32.5728	33	23

Page 15 of 16

Page 16 of 16

EV Table 2.doc

Example of the summary output of AnalyseDNA.m program
(summary for 10 3 by 3 montage images)

1	1107	163.312	79.3918	1.59819	0.398738	0.724461	0.137996	14.0612	3.3315	0.903207	0.0310365	0.701248	0.073176	6449.26	3186.95	41.339	18.332	42.9416	10.3393
2	22.4103	14.334	30.3906	22.4584	0.399942	0.727935	0.138162	14.2634	3.43248	0.904371	0.0311453	0.701377	0.0730219	6784.37	3895.1	42.2416	17.0945	43.8219	17.393
3	33.0245	33.3162	31.3609	21.7507	0.397933	0.725532	0.131153	14.0853	3.33904	0.903247	0.0314182	0.702891	0.0720005	6881.08	3575.81	44.7167	20.5918	46.3936	21.2287
4	24.8313	16.2062	36.6117	25.1765	0.389642	0.727142	0.142518	14.3331	3.43812	0.902766	0.03171254	0.693545	0.0733889	6997.68	4212.87	43.1198	19.8382	44.5361	20.5112
5	33.7349	35.9802	32.505	24.193	0.400492	0.721048	0.141204	14.338	3.45065	0.904132	0.0319806	0.70023	0.0736884	7050.22	4163.04	44.0559	21.4761	45.8322	22.1282
6	34.3315	16.4359	33.6367	26.3398	0.425512	0.724414	0.133721	14.0974	3.35686	0.904614	0.0327731	0.696204	0.0782455	6843.2	3924.32	44.266	19.3396	45.9405	19.6324
7	24.5144	15.2164	33.8055	23.4472	0.441022	0.694813	0.185667	11.728	5.27441	0.893711	0.0401739	0.704526	0.0832393	5162.51	4393.31	34.9703	21.1869	36.16	22.0024
8	24.8301	17.0423	43.0689	25.4082	0.405147	0.723121	0.137599	14.3833	3.40593	0.904384	0.0331179	0.701374	0.0731602	6645.87	3767.48	42.2752	17.2733	43.9791	17.7824
9	22.0402	33.3194	31.3941	20.355	0.401986	0.72568	0.136518	14.1402	3.43794	0.905208	0.0336286	0.700331	0.0746473	6574.54	4027.38	41.6064	17.8994	43.143	18.2489
10	24.2469	35.8119	30.5974	22.3802	0.400372	0.717147	0.139045	13.8205	3.48556	0.904275	0.0371911	0.702759	0.07372	6587.18	3788.39	44.2141	20.7694	45.8765	21.136
	24.4382	16.3159	33.7938	25.4094															

CLAIMS

What is claimed is:

1. A method of predicting a property of a manipulation of cells based
5 upon a descriptor, said method comprising:
 providing a plurality of cells;
 manipulating said plurality of cells;
 capturing a morphological value from said plurality of cells;
 assigning a degree of presence of said morphological value; and
10 storing said morphological value and said degree of presence;
 wherein said descriptor is derived from a first component of a cell and
a second component of said cell, said capturing said morphometric value from said
plurality of cells comprises determining a relationship between said first component
and said second component.
- 15 2. The method of claim 1 wherein said first component and said second
component are selected from a protein, a protein modification, a nucleic acid, a lipid,
a carbohydrate, a subcellular structure and an organelle.
3. The method of 1 wherein said step of manipulation occurs in a manner
selected from a electrical source, a chemical source, a thermal source, a gravitational
20 source, a nuclear source, a temporal source, and a biological source
4. The method of claim 3 wherein said chemical source is selected from a
pharmacological candidate and a drug screening library.
5. The method of claim 1 wherein said morphological value is selected
from a count, an area, a perimeter, a length, a breadth, a fiber length, a fiber breadth, a
25 shape factor, an elliptical form factor, an inner radius, an outer radius, a mean radius,
an equivalent radius, an equivalent sphere volume, an equivalent prolate volume, an
equivalent oblate volume, an equivalent sphere surface area, an average gray value, a
total gray value, and an optical density.
6. The method of claim 1 wherein said degree of presence is
30 multiple of a quantized value.

7. A computer program product for populating a database with manipulated biological information, said computer program product comprising:
- code for providing a plurality of cells in various stages of the cell cycle, said stages of the cell cycle including at least one selected from interphase, G0 phase, G1 phase, S phase, G2 phase, M phase, prophase, prometaphase, metaphase, anaphase, and telophase;
 - code for manipulating said cells in said various stages of cell cycle development to form a plurality of manipulated cells;
 - code for capturing an image of said plurality of manipulated cells;
 - code for determining a descriptor from said image for said manipulated cells;
 - code for populating a database with said descriptor;
 - wherein said image includes a first component of a cell and a second component of said cell; and
 - a computer readable storage medium for holding the codes.
8. The computer program product of claim 7 wherein said first component and said second component are selected from a protein, a protein modification, a nucleic acid, a lipid, a carbohydrate, a sub-cellular structure and an organelle.
9. The computer program product of claim 7 wherein said image is a digitized representation of said plurality of manipulated cells.
11. The computer program product of claim 9 wherein said digitized representation provides a density value of said plurality of manipulated cells.
11. The computer program product of claim 7 wherein said descriptors comprise numeric or logical values.
12. The computer program product of claim 11 wherein said values further comprises a nucleotide.
13. The computer program product of claim 11 wherein said values further comprises an amino acid letter.
14. A system for capturing images of cells or cell structures, the system comprising:
- a cell holder comprising a plurality of sites in a spatial orientation, each of the sites being capable of holding a plurality of cells to be imaged;

an image capturing device coupled to the cell holder, the image capture device being adapted to capture at least one image in at least one of the plurality of sites;

an illumination apparatus comprising a liquid light guide coupled to the plate for highlighting the plurality of cells in a relatively even spatial manner for image capturing purposes;

an image processing device coupled to the image capturing device, the image capturing device being adapted to convert the image into a digital representation; and

a database storage device comprising a database management element coupled to the image capturing device, the database storage device being adapted to retrieve the digital representation of the image from the image processing device and storing the digital representation.

15. The system of claim 14 further comprising a stage comprising a device for moving the cell holder in a spatial direction to traverse across the cell holder in the spatial orientation.

16. The system of claim 14 wherein the illumination apparatus comprises sub-elements, at least one of the sub-elements being positioned away from the image capturing device to prevent a possibility of vibration from the one sub-elements to be transmitted to the image capturing device.

17. The system of claim 14 wherein the digital representation comprises a plurality of regions and objects.

18. The system of claim 14 further comprising a computing device connected between the database storage device and the image processing device.

19. The system of claim 14 wherein the image capturing device comprises a magnification of at least 1X and greater to capture the image of the site.

20. The system of claim 14 wherein the plurality of sites comprises at least 96 sites.

21. The system of claim 14 wherein the liquid light guide characterized as a flexible member that substantially prevents vibration from the an element of the illumination apparatus to be transferred to the image capturing device.

22. The system of claim 14 wherein the spatial direction can be selected from an x-direction, a y-direction, or a z-direction in a Cartesian coordinate system.

23. The system of claim 14 wherein the each of the sites comprises
5 a volume that is sufficient to prevent a solution therein from evaporating in a substantial manner that may influence the image capturing.

24. A method for identifying a mechanism of action for a first compound, the method comprising the steps of:

receiving the first compound;
10 measuring at least one feature of a cellular phenotype to define a target phenotype;
identifying additional compounds providing a feature similar to the feature identified in the measuring step; and
characterizing the first compound in terms of distance from a specific
15 target phenotype having known characteristics.

25. The method of claim 24 comprising the further step of storing the additional compounds and their associated phenotypes in a database.

26. A method for identifying a mechanism of action for a cellular stimulus, the method comprising the steps of:
20 receiving cells of interest;
measuring at least one feature of the cells to define and quantify a target phenotype;
identifying additional compounds providing a feature similar to the feature identified in the measuring step; and
25 characterizing the first compound in terms of distance from a specific target phenotype having known characteristics.

27. A method for identifying information relevant to at least one of a mechanism of action and cellular activity by utilizing assay data to elucidate a phenotype, the method comprising the steps of:

30 identifying a target protein;
identifying positive and negative biochemical hits related to the target protein;
defining the target phenotype utilizing the positive and negative hits;

identifying other compounds providing similar features; and
characterizing the first compound in terms of distance from a specific
target phenotype having known characteristics.

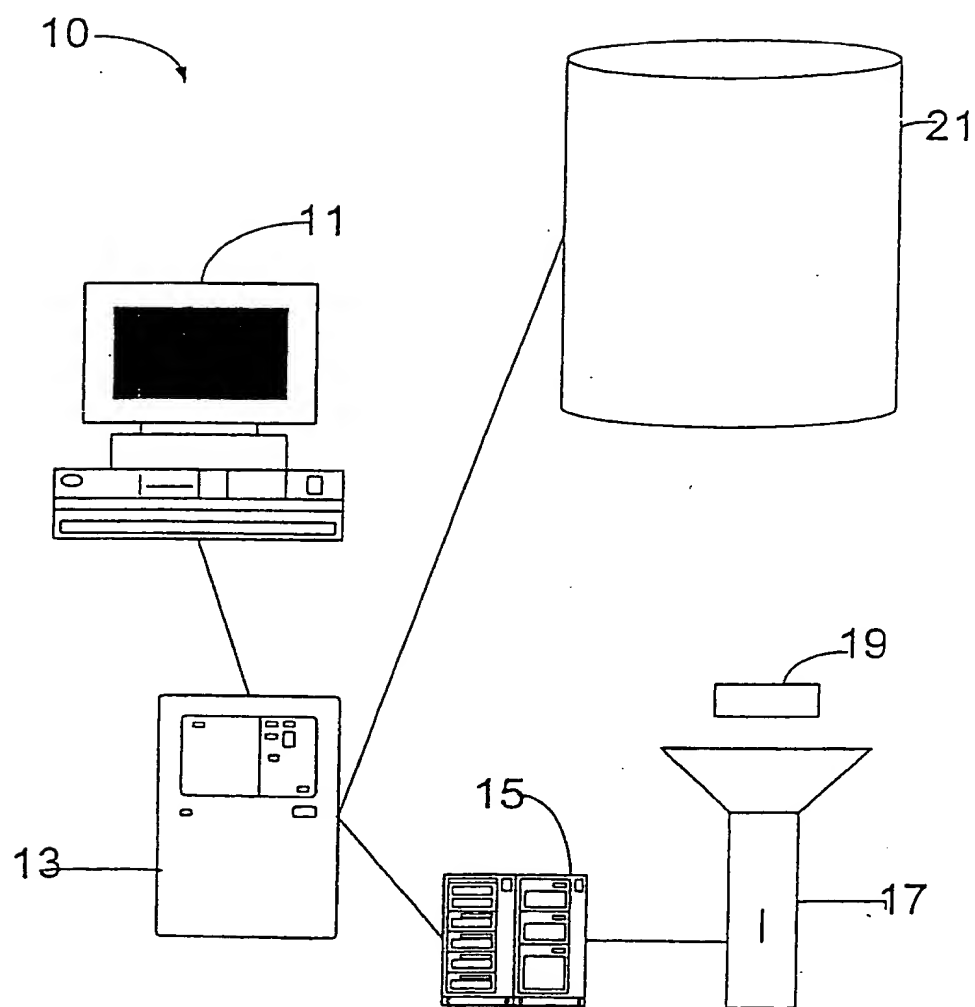


FIG. 1

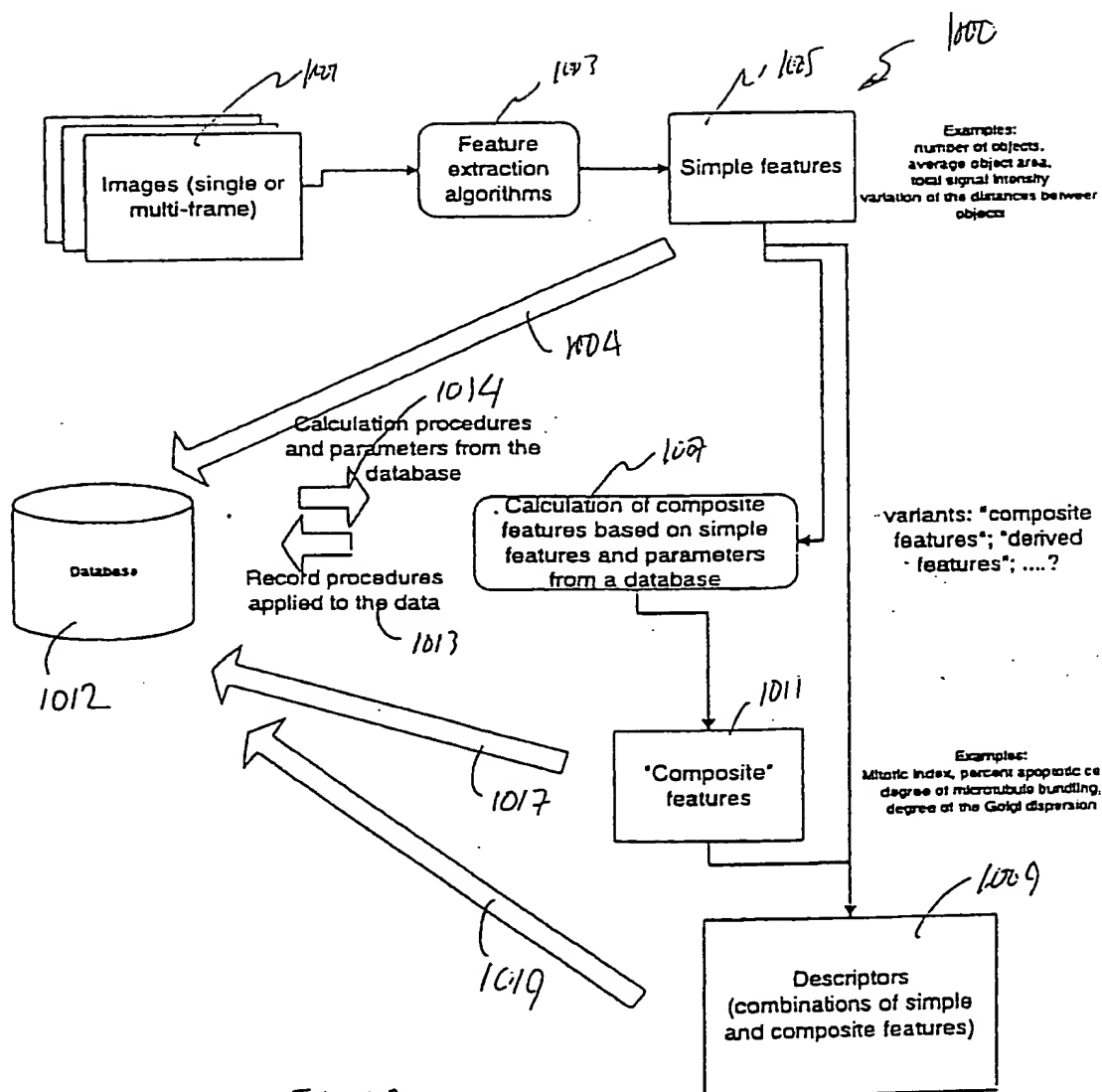


FIG. 1A

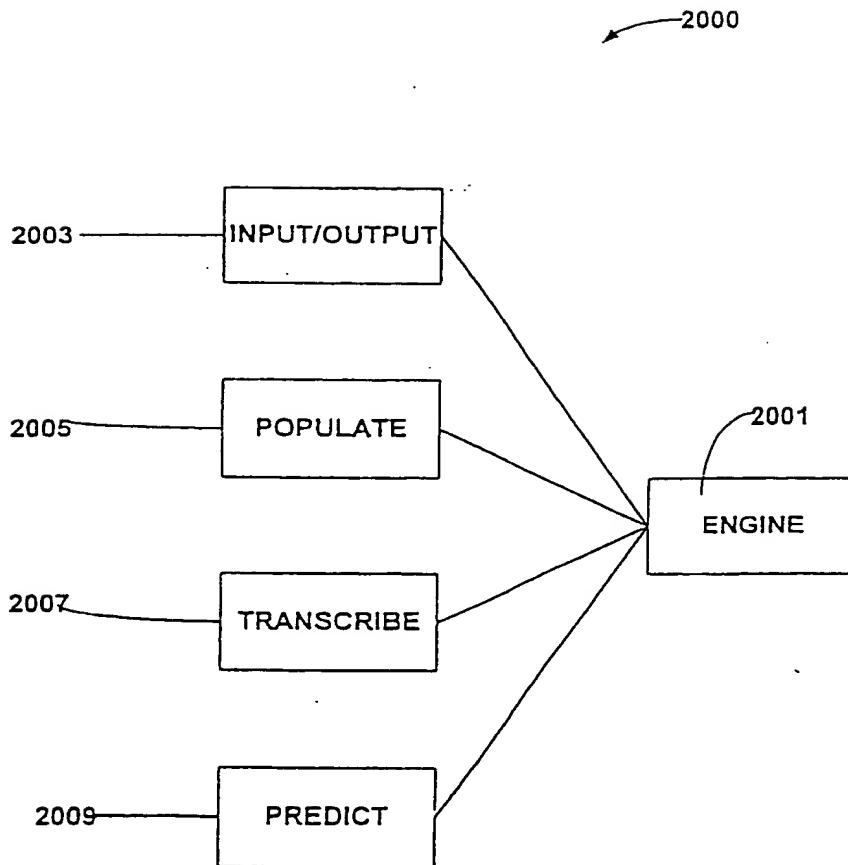


FIG. 1B

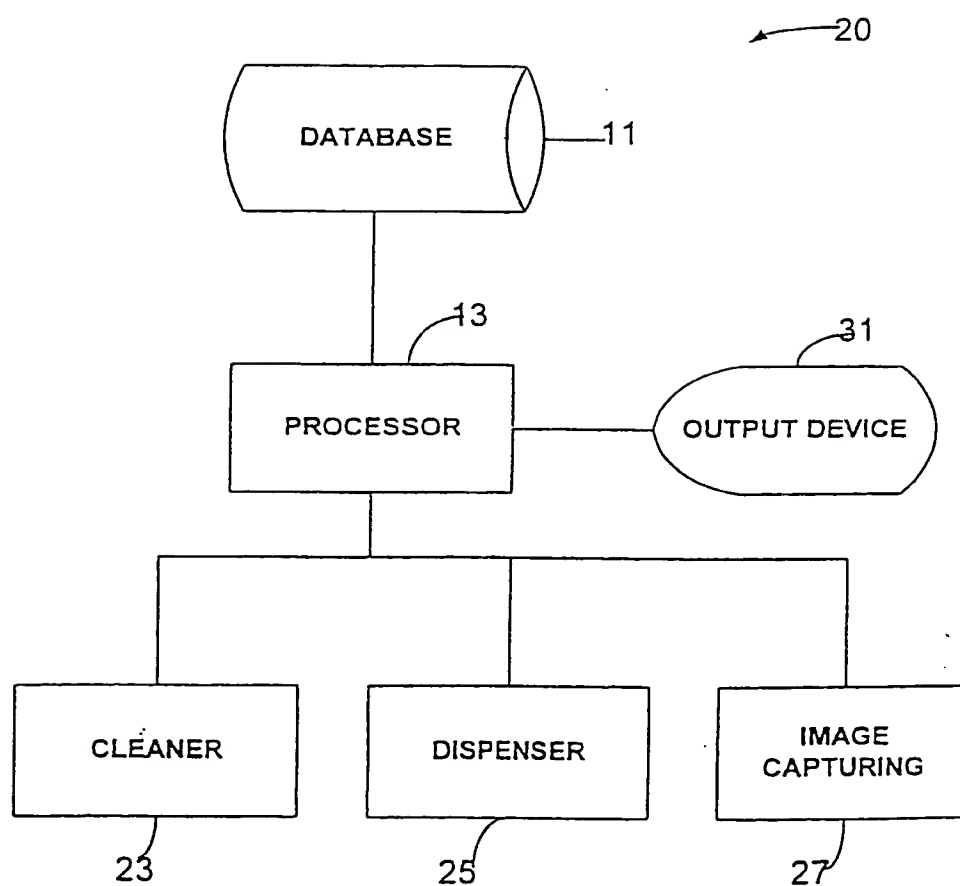
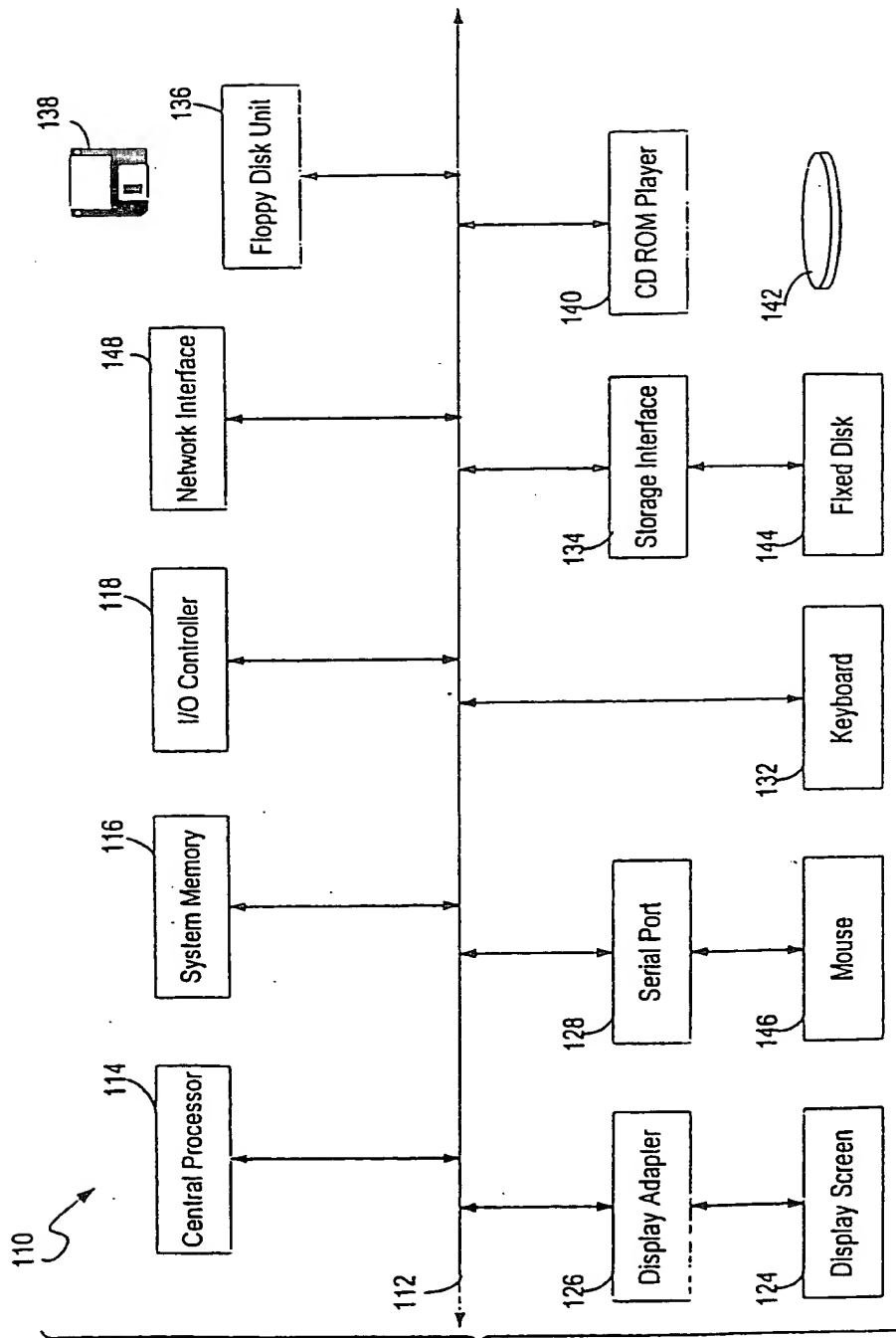


FIG. 2



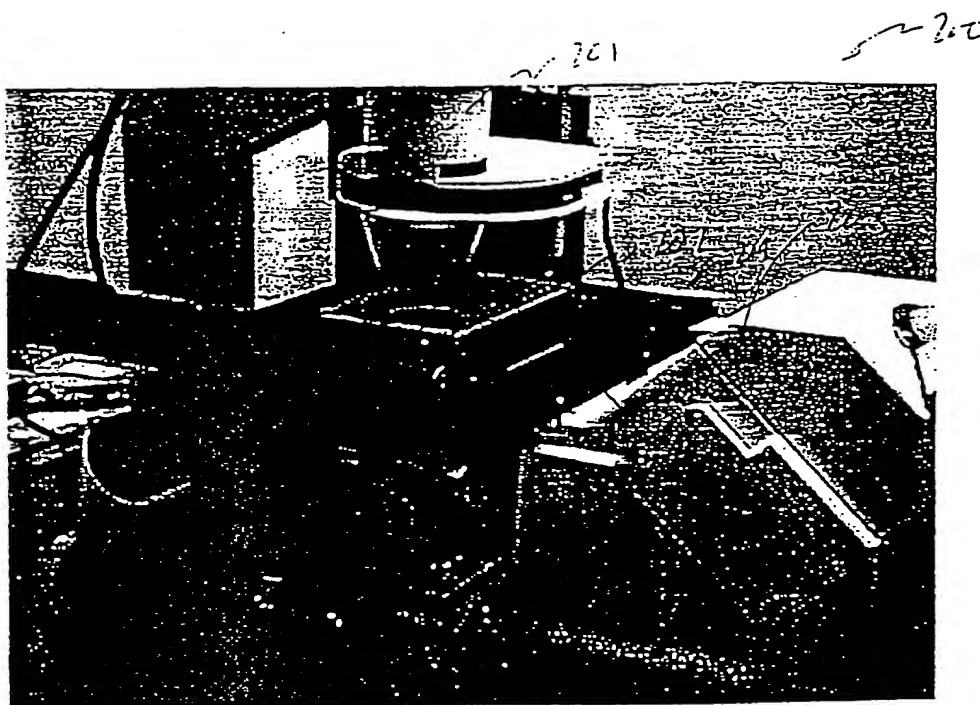


FIG. 4

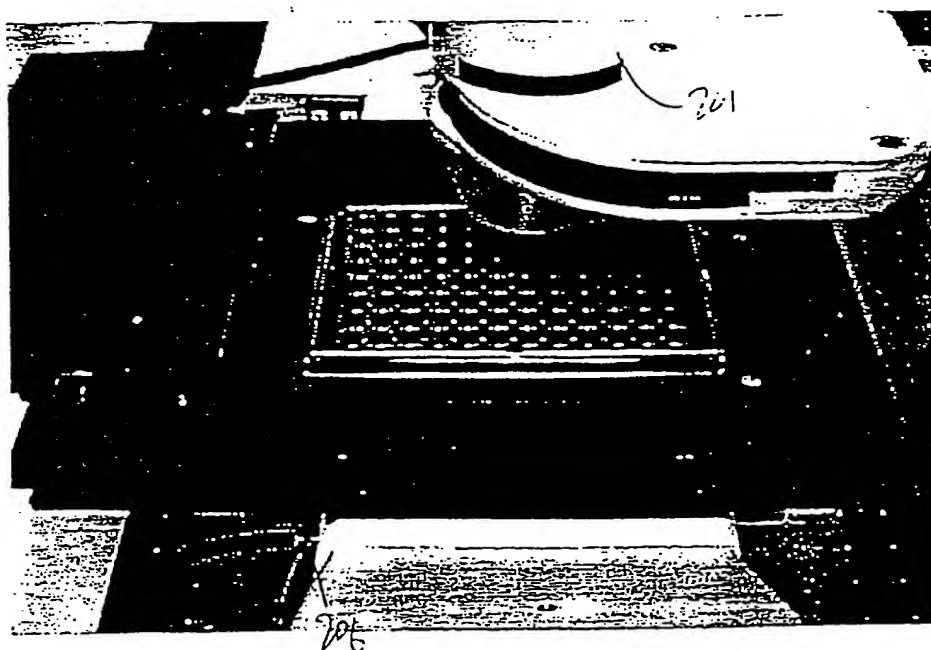
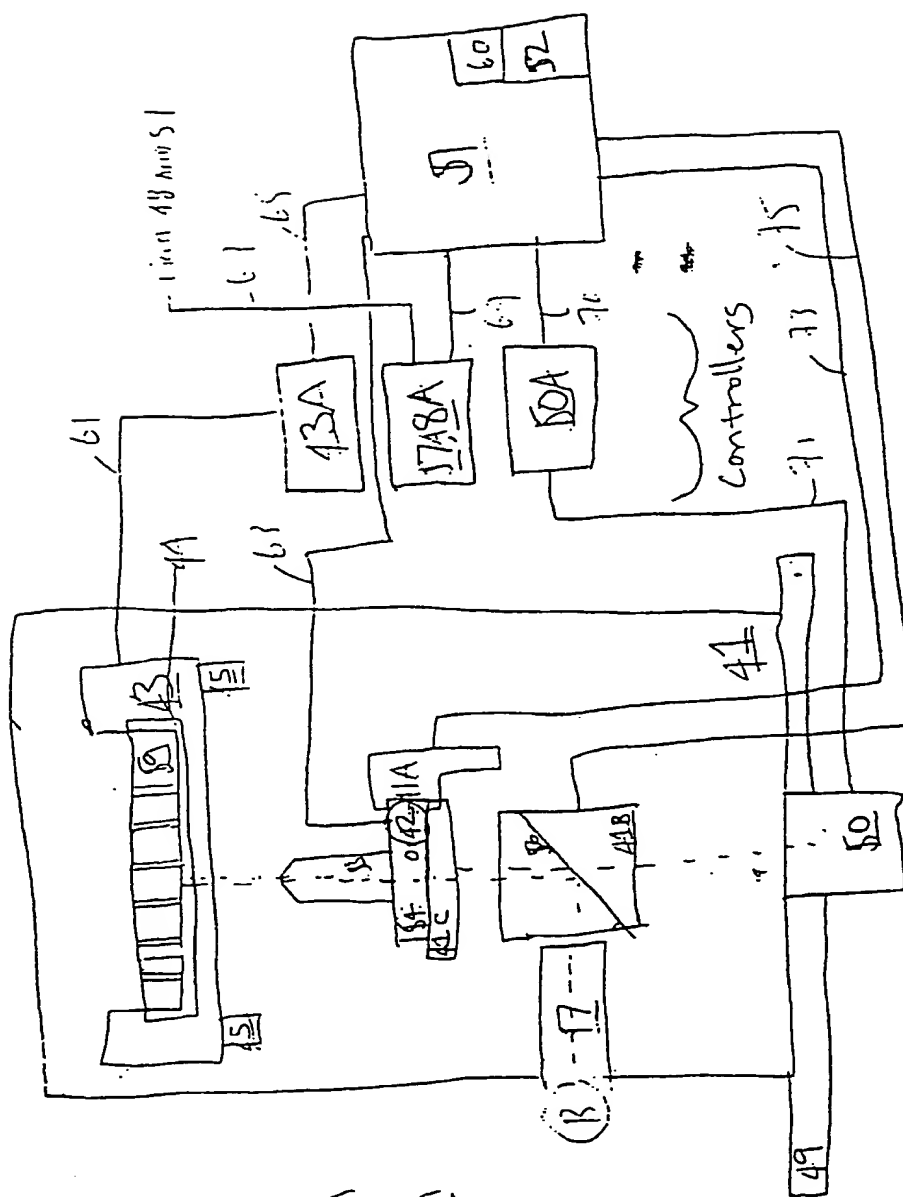


FIG 5



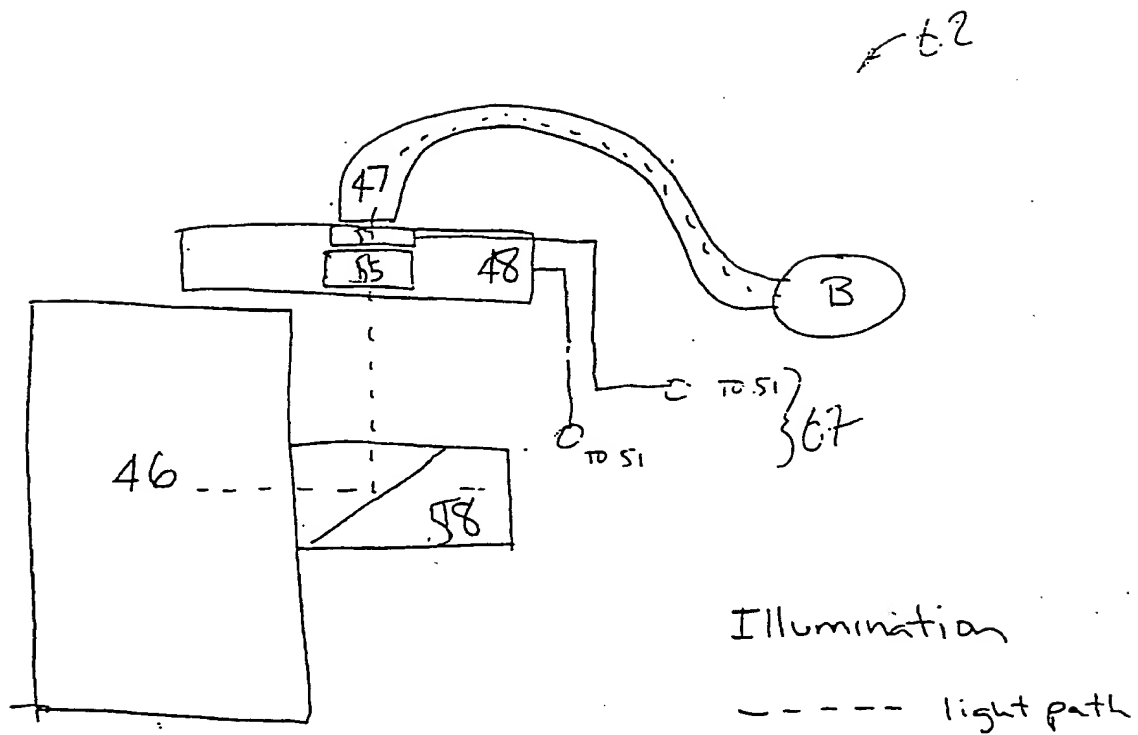


FIG. 5B

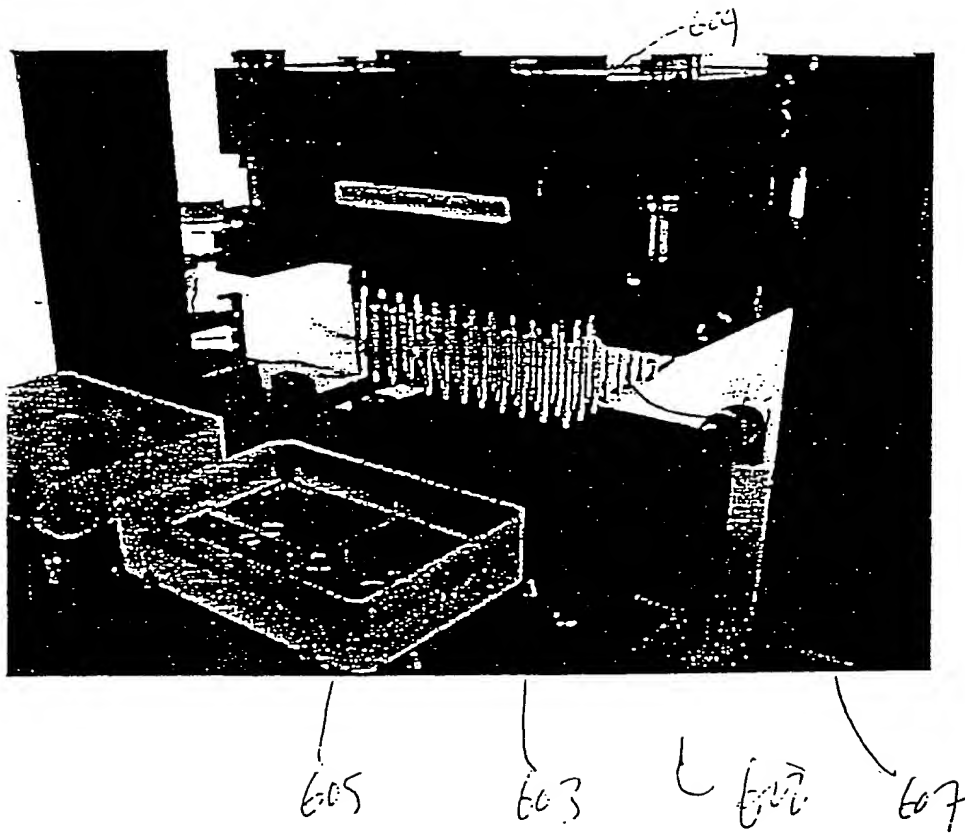


FIG. 6

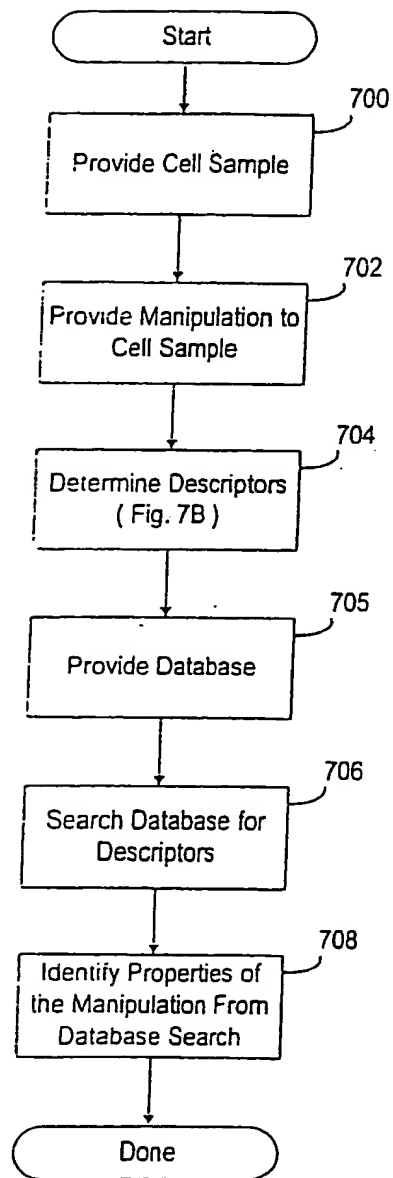


Fig. 7A

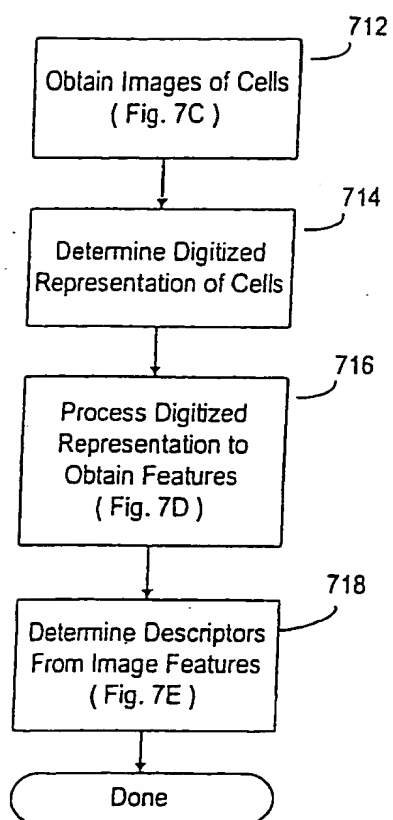


Fig. 7B
Step 704 of Fig. 7A

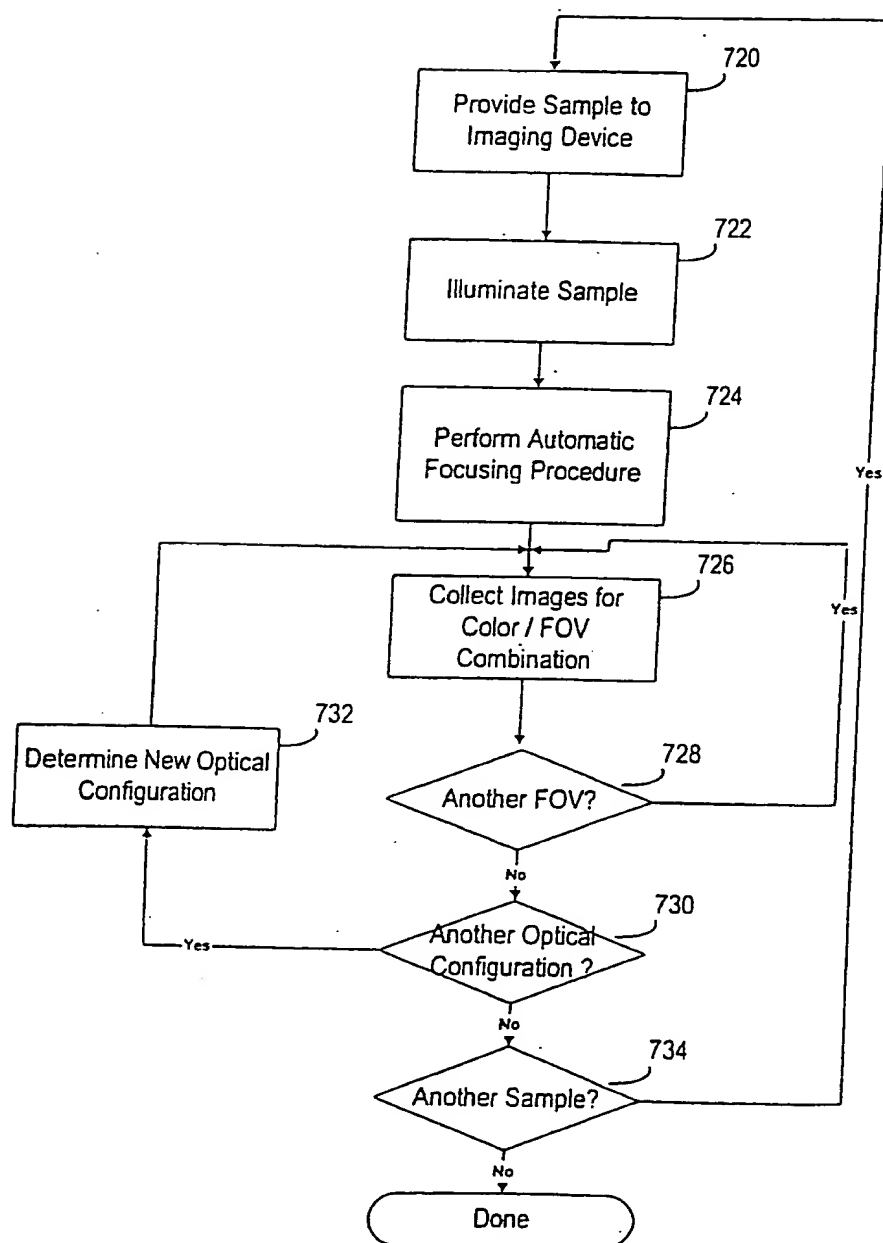


Fig. 7C
Step 714 of Fig. 7B

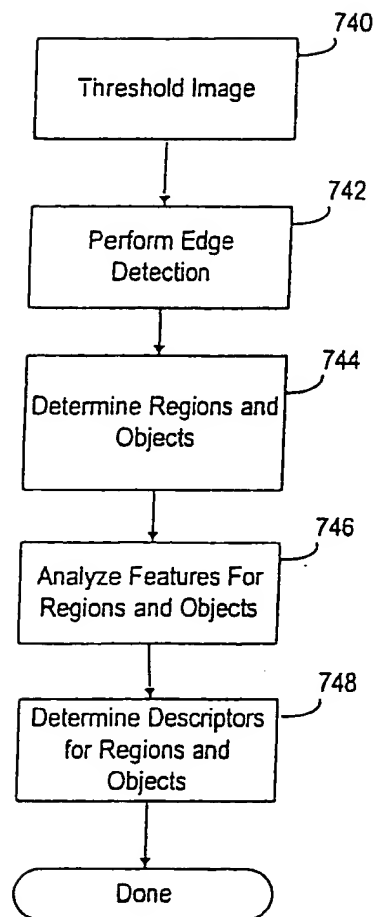


Fig. 7D
Step 716 of Fig. 7B

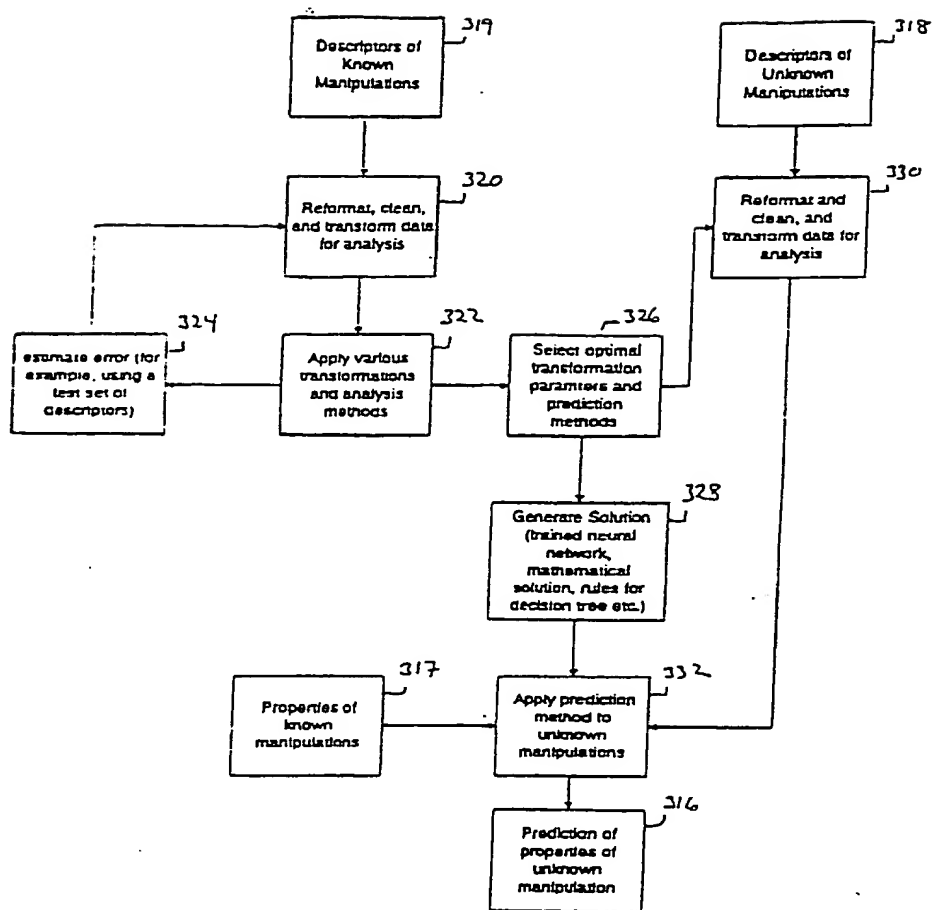


FIG. 7E

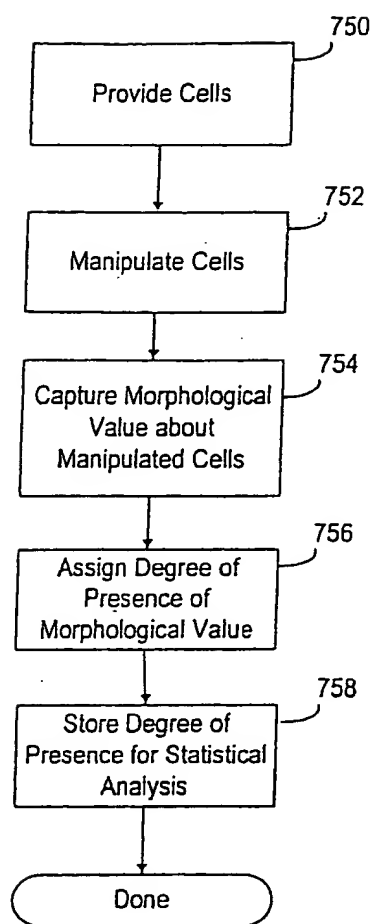


Fig. 7F

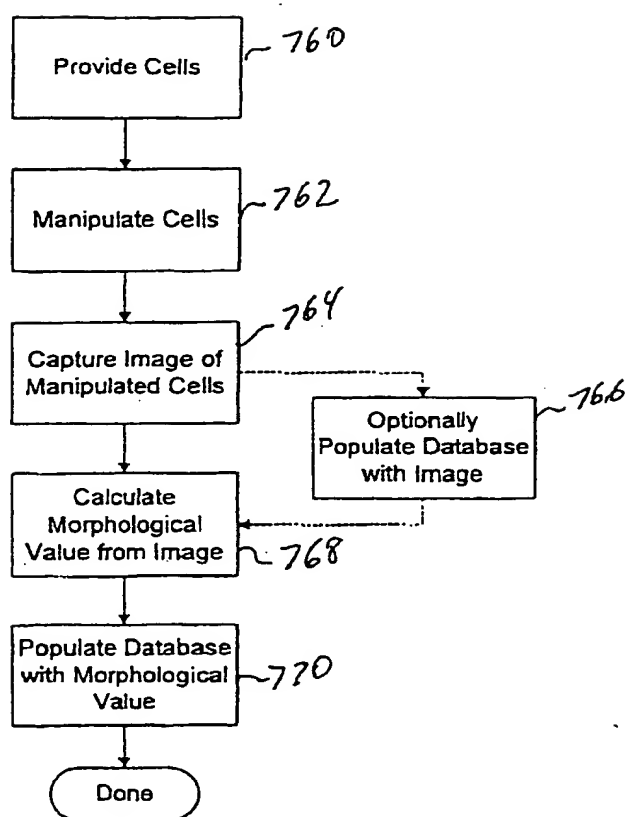


Fig 7G

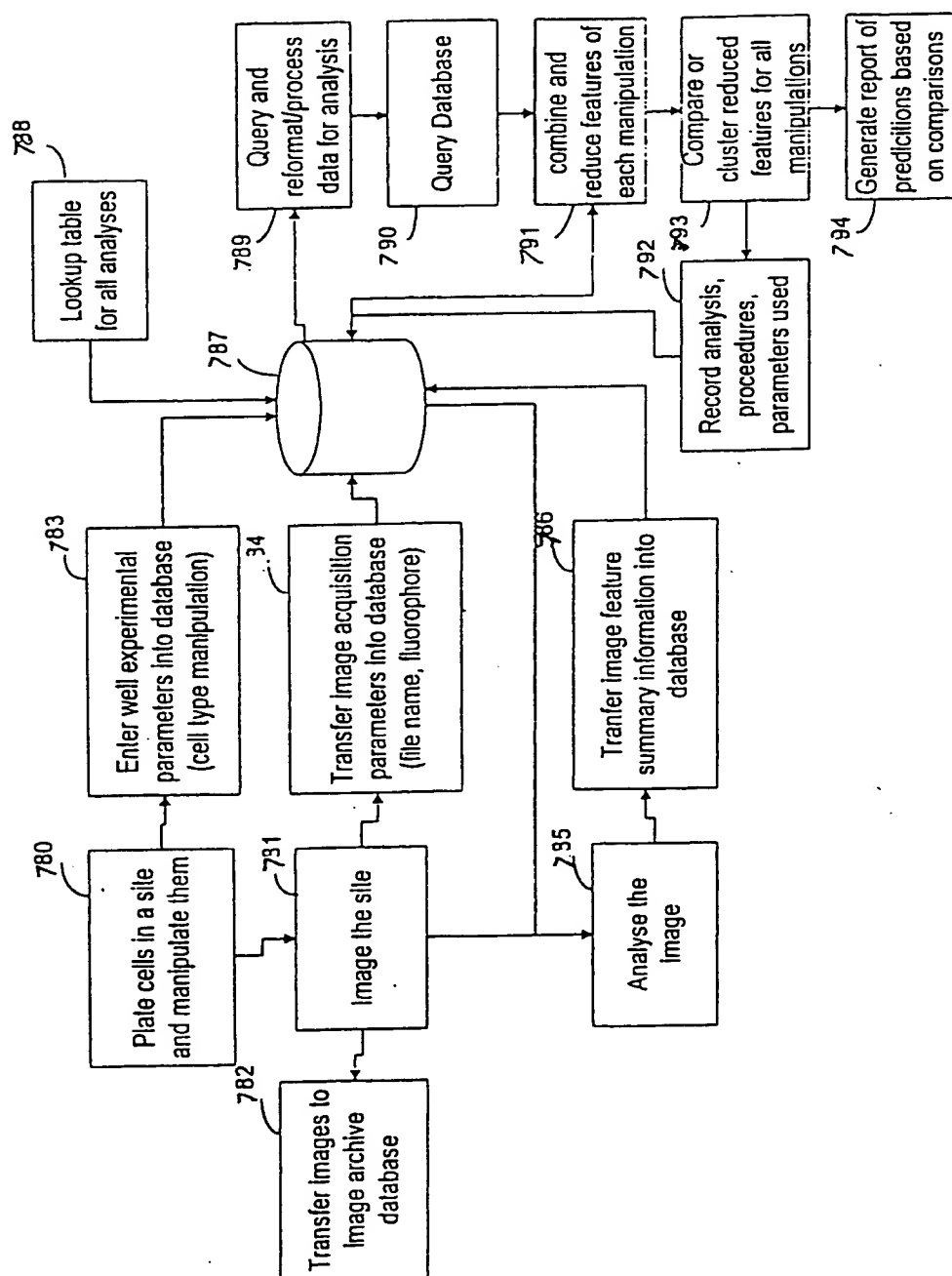


Fig. 7 H

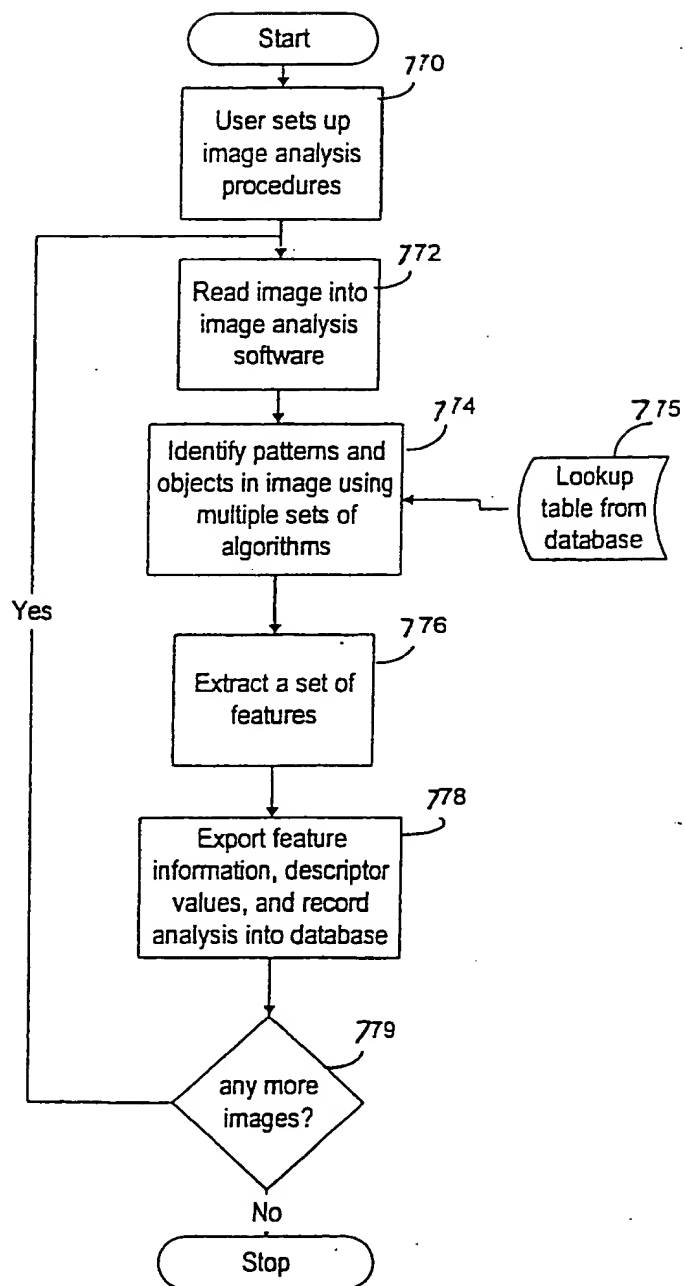


Fig. 71

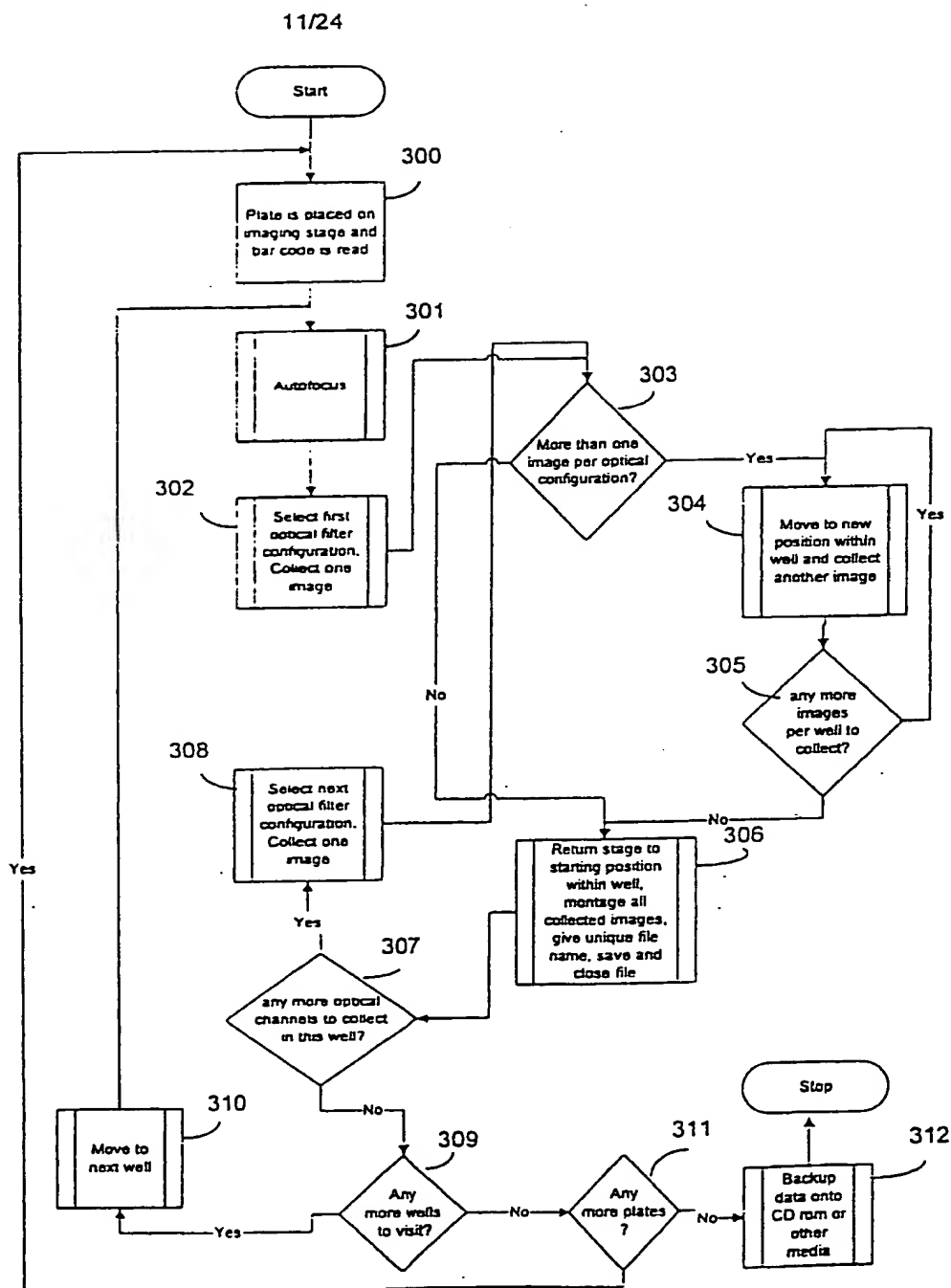


Fig. 7J

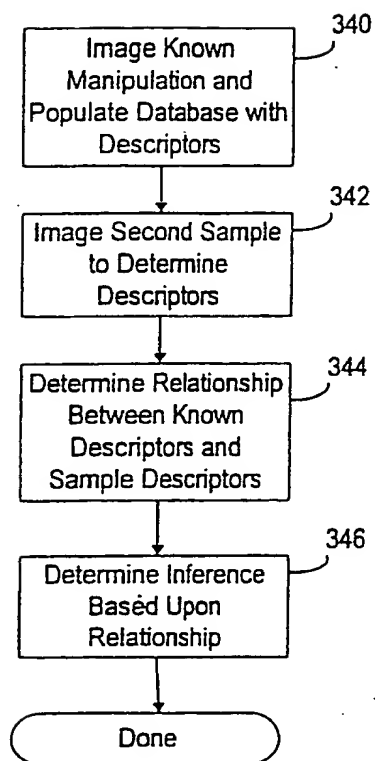


Fig. 7K

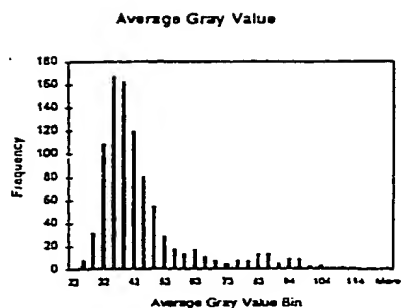


Fig. 8A

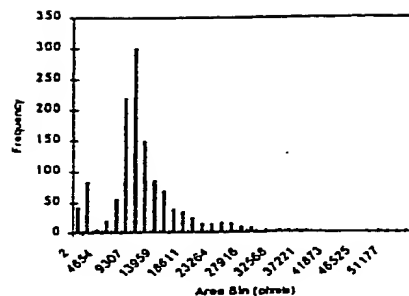


Fig. 8B

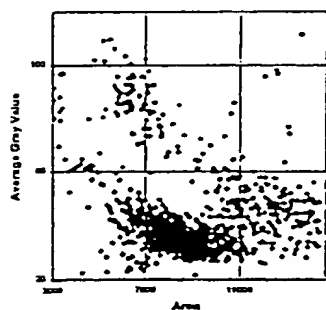


Fig. 8C

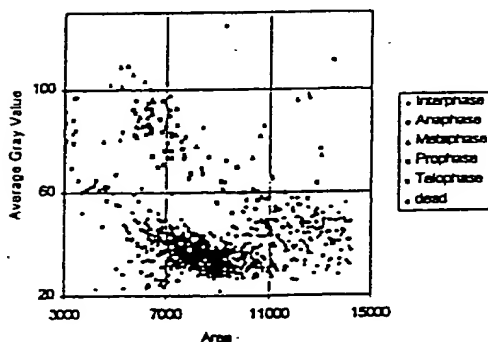


Fig. 8D

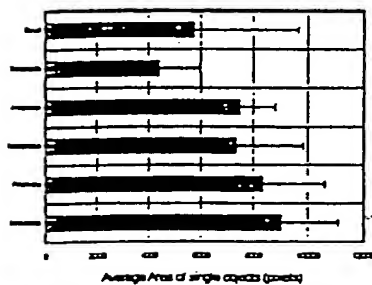


Fig. 8E

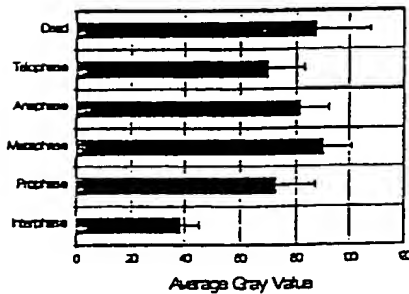


Fig. 8F

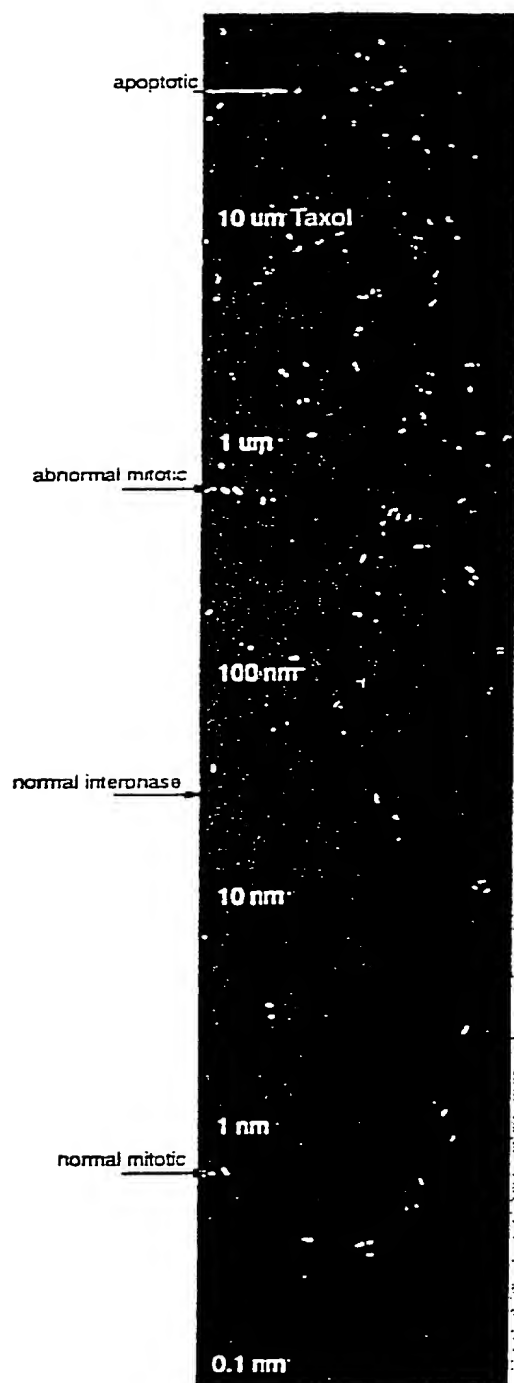


Fig. 9

MDCK cells treated with Taxol for 4.5 hours

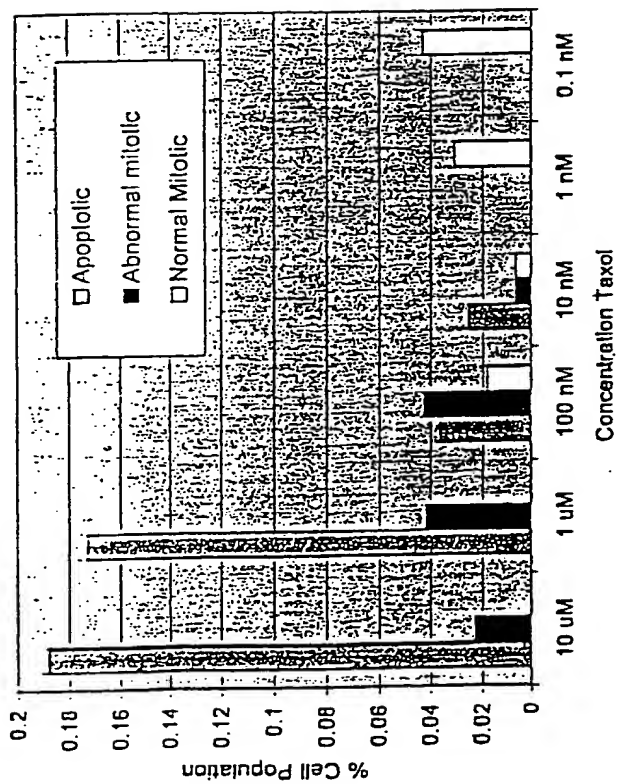
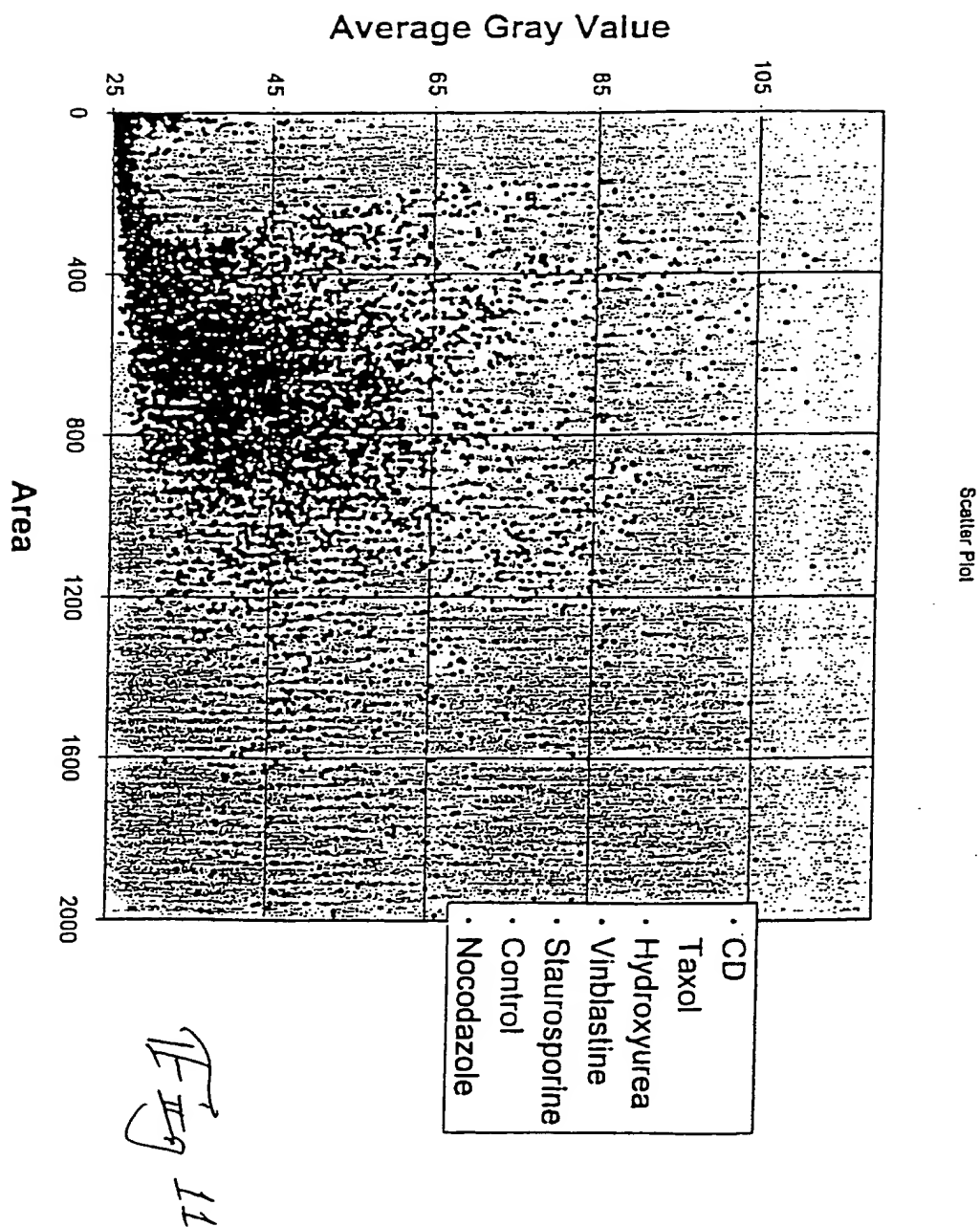
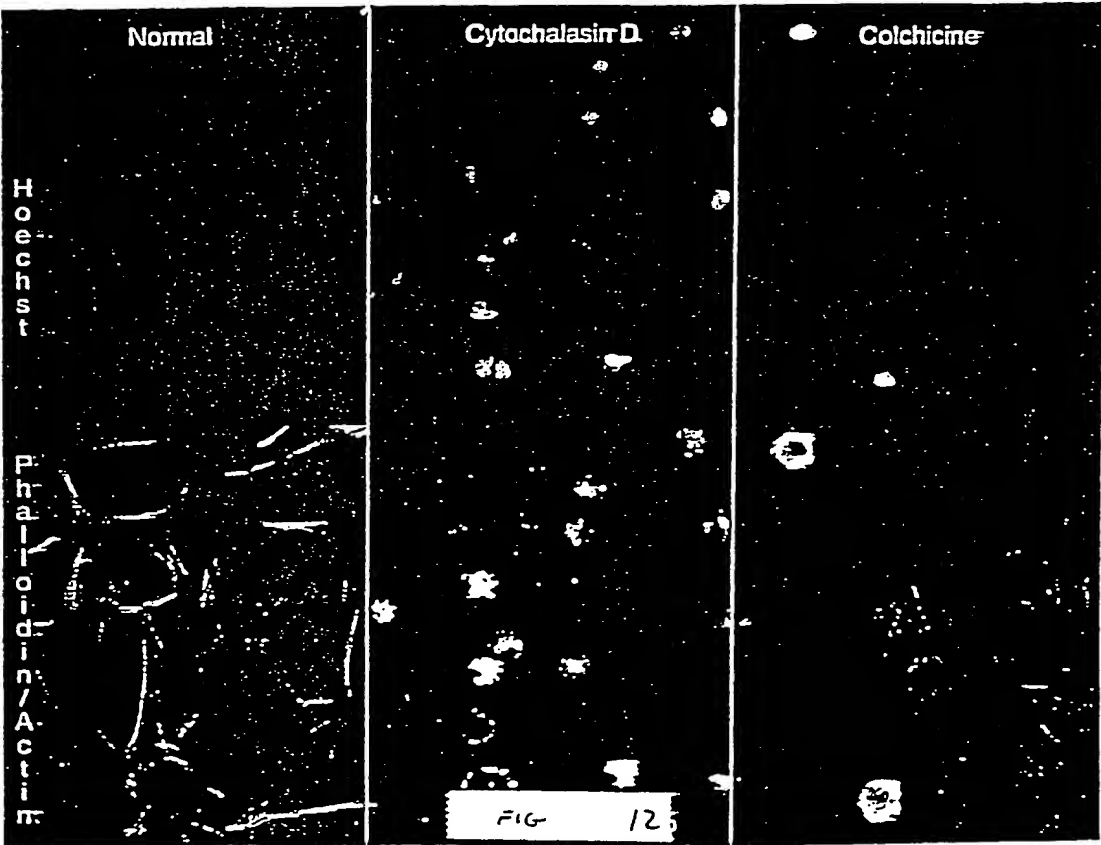


Fig 10





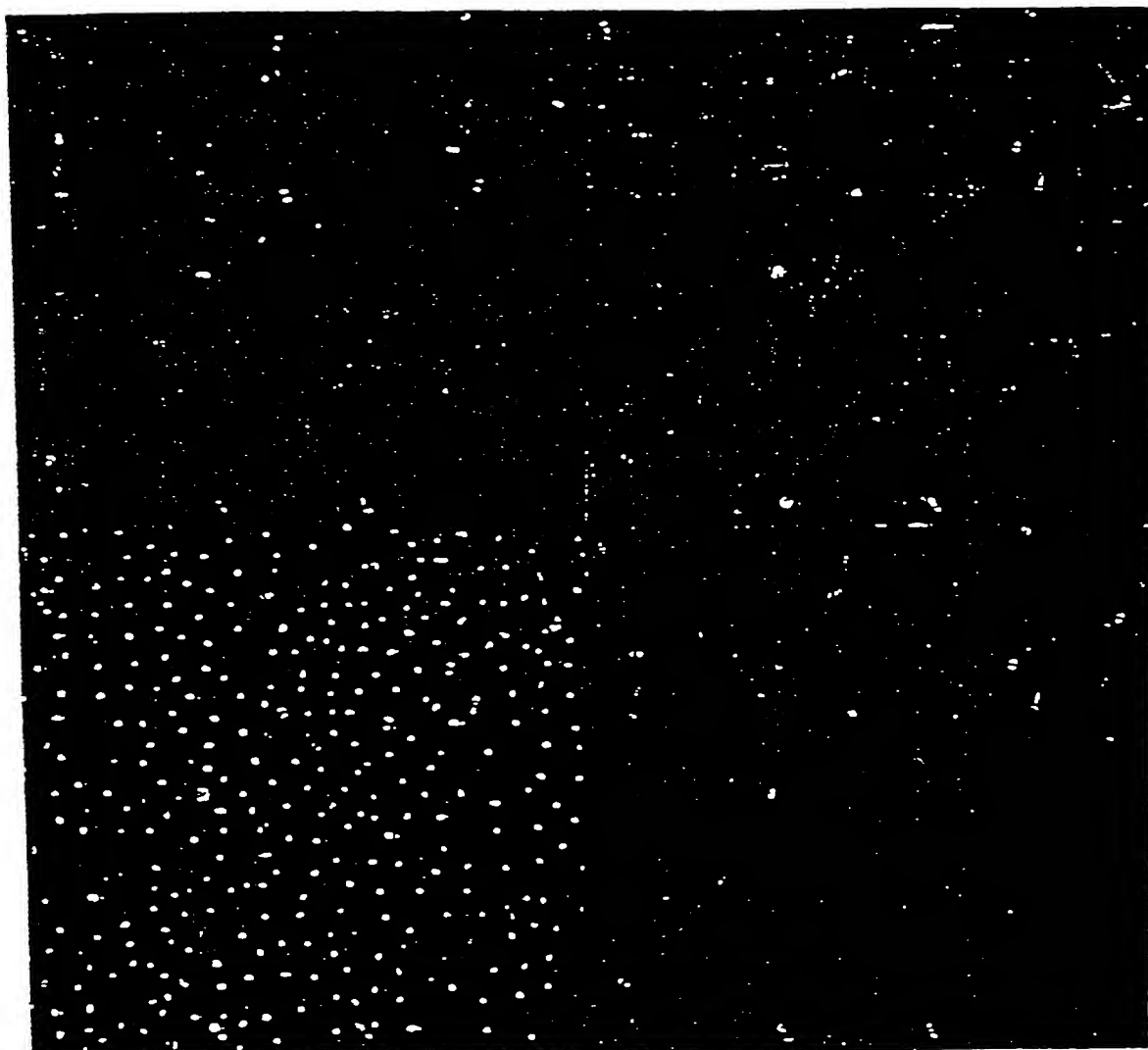
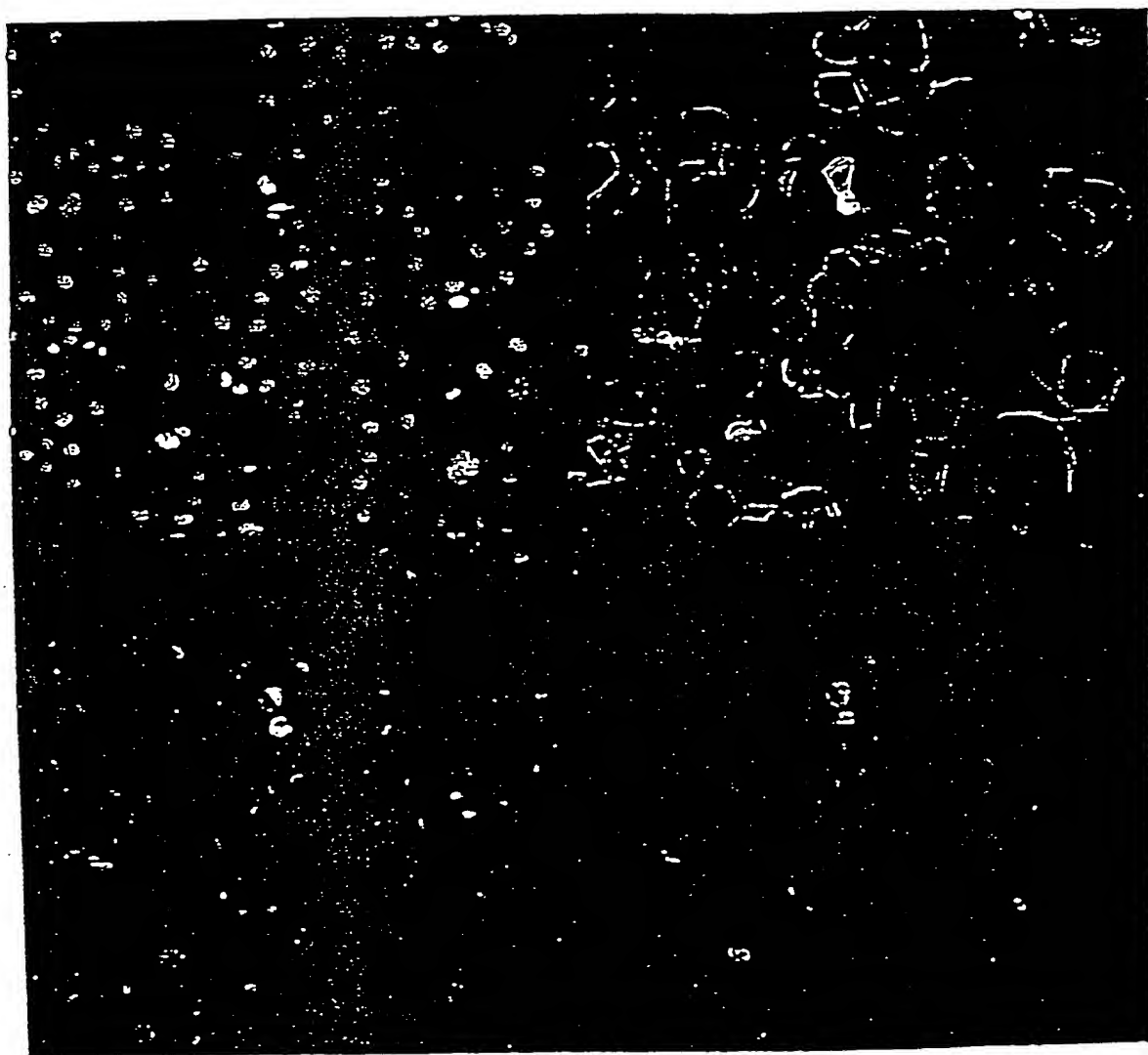


Fig 13



Page 14

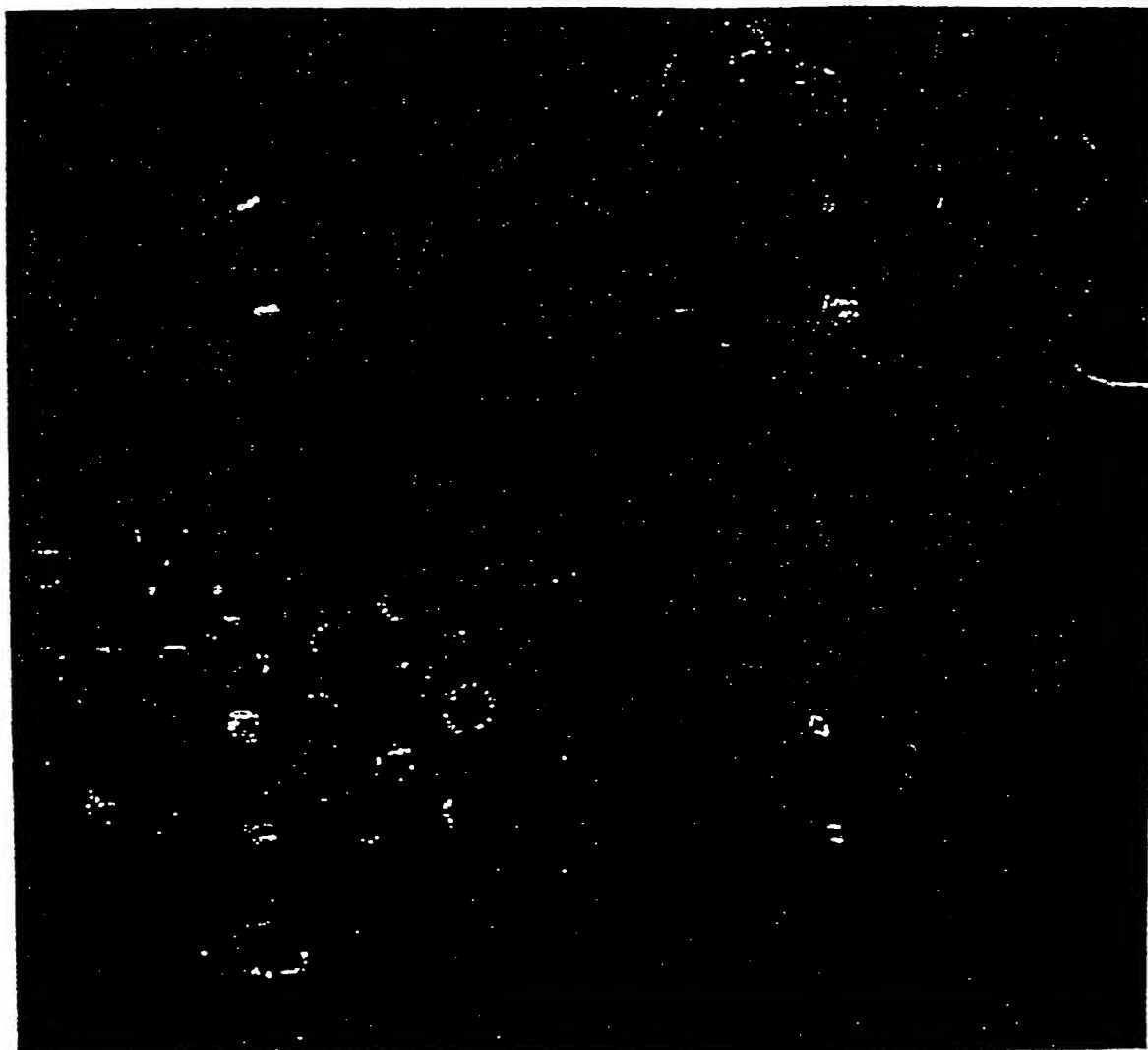


Fig 15

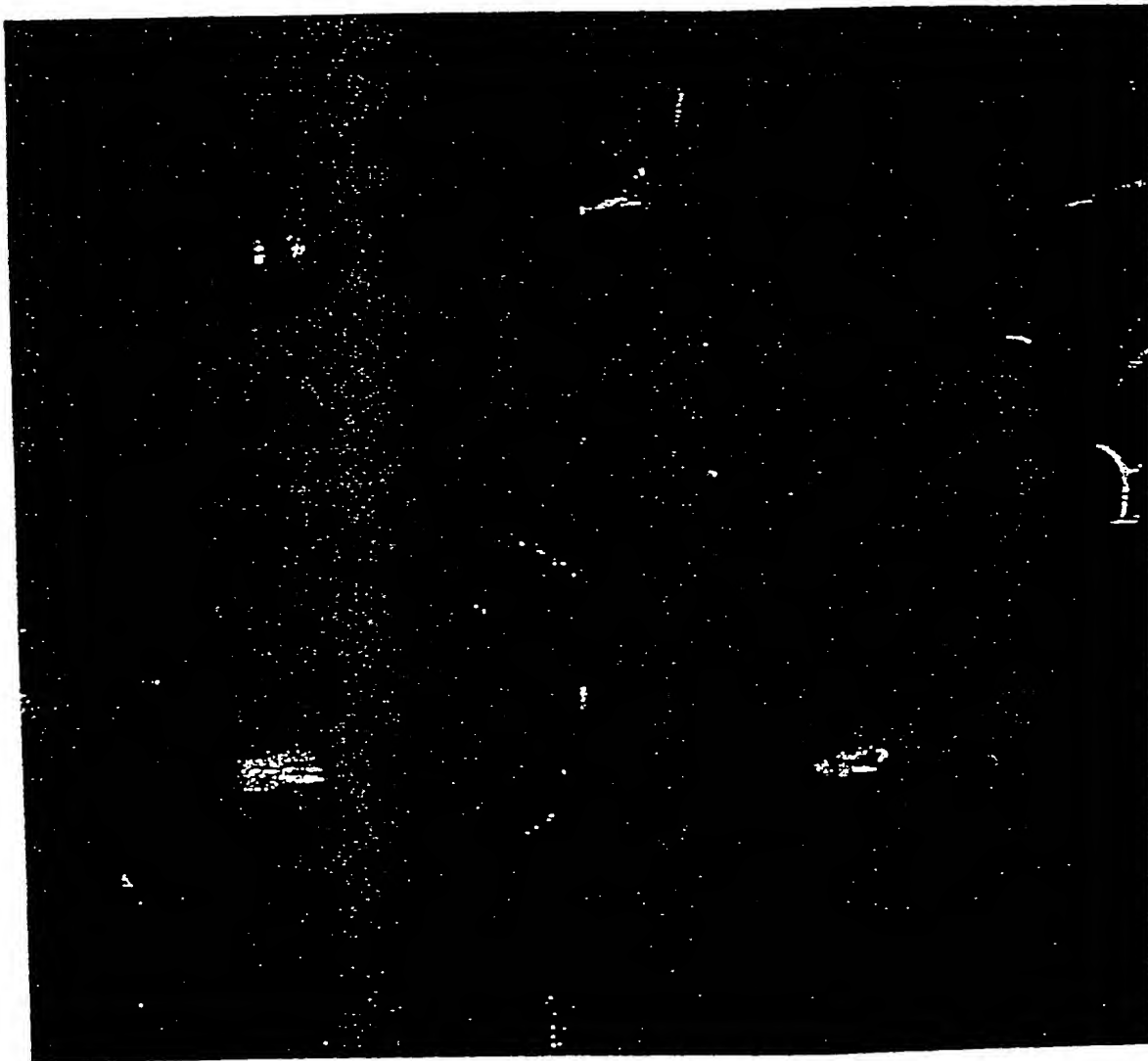


Fig 16



Fig 17

Conversion of morphometric parameters into nucleic acid code
and clustering of the resulting sequences using Neighbor
Joining method.

Compound:	Measurements																							
	Count	Area	Perimeter	Length	Breadth	Fiber length	Fiber breadth	Shape factor	Ell. form factor	Inner radius	Outer radius	Mean radius	Equiv. radius	Equiv. sphere vol.	Equiv. prolate vol.	Equiv. oblate vol.	Equiv. sphere surface area	Average gray value	Total gray value	Optical density	Radial dispersion	Texture Difference Moment	EFA Harmonic 2, Semi-Maj	EFA Harmonic 2, Semi-Min
Control	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	a	t	t	
Taxol	a	t	t	t	t	t	t	t	a	t	t	t	t	t	t	t	t	t	t	t	t	t	t	
CD	c	a	a	a	t	a	t	t	c	a	a	a	a	a	a	a	a	t	a	a	a	t	a	g
Nocodazol	c	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	
Staurosporine	g	g	c	a	a	t	a	a	t	g	a	a	a	t	g	g	g	a	a	t	a	t	a	a
Vinblastine	c	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	g	t	t	t	t	t	
Hydroxyurea	g	t	t	t	t	t	t	g	t	t	t	t	t	t	t	t	t	t	t	c	t	a	t	t

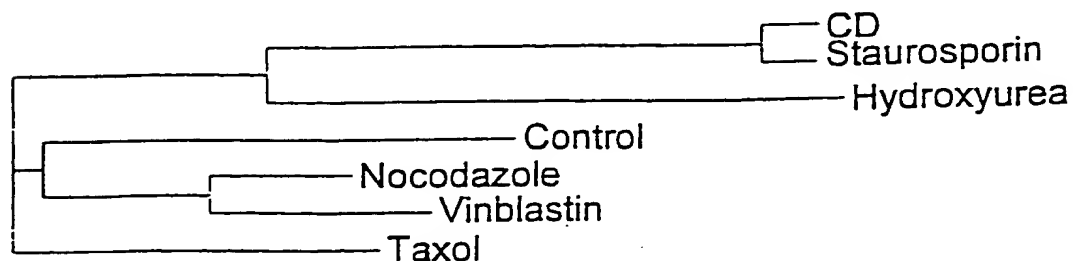


Fig 18

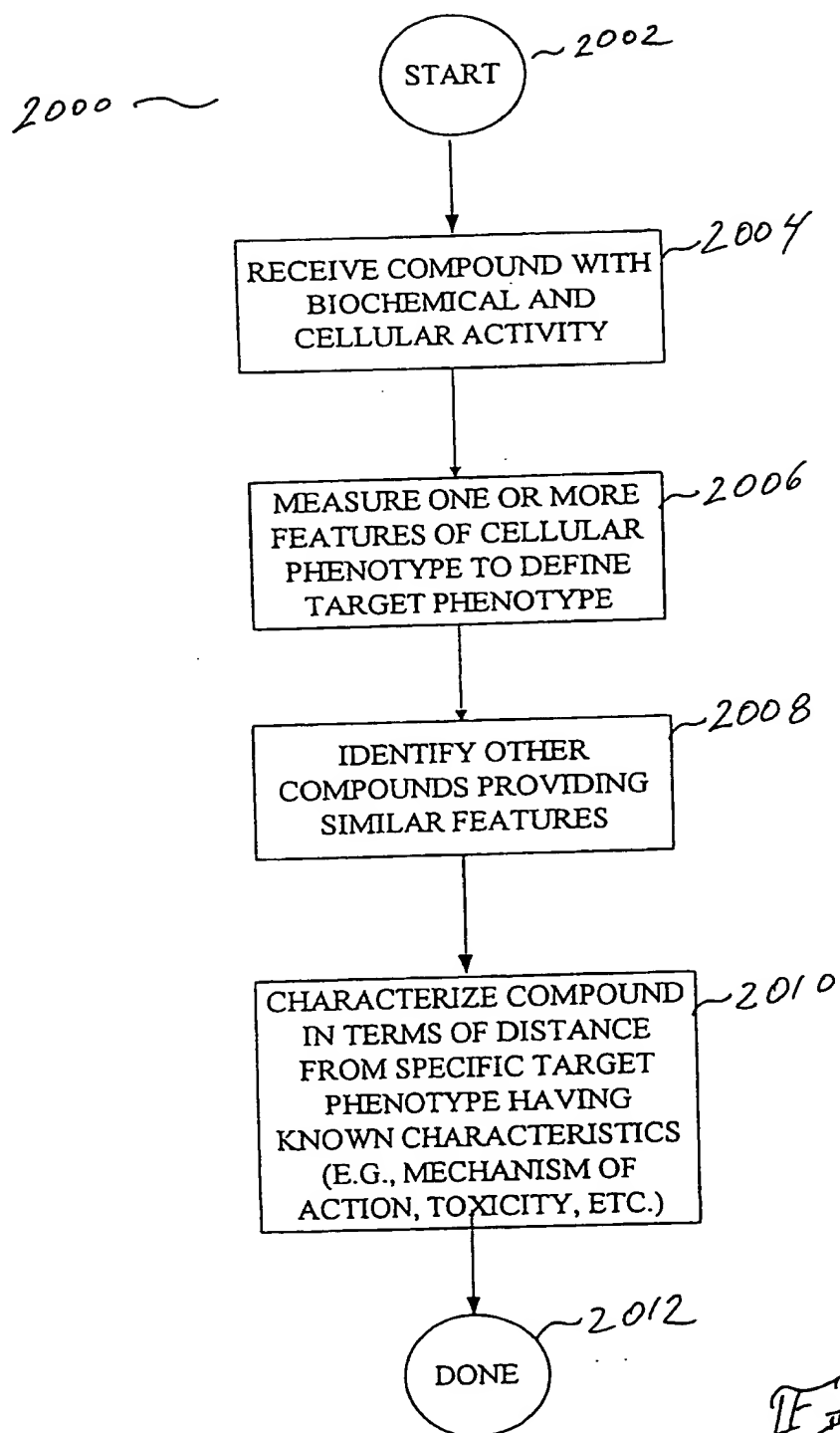


Fig 20

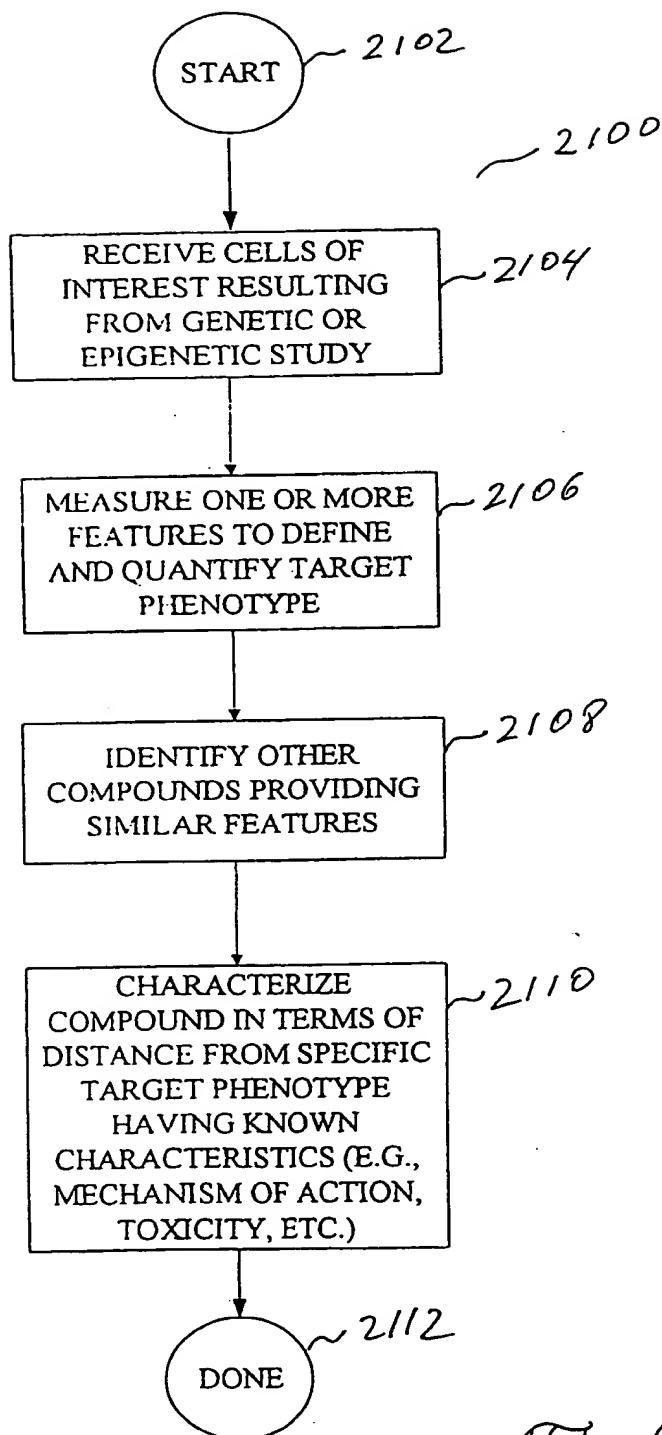


Fig 21

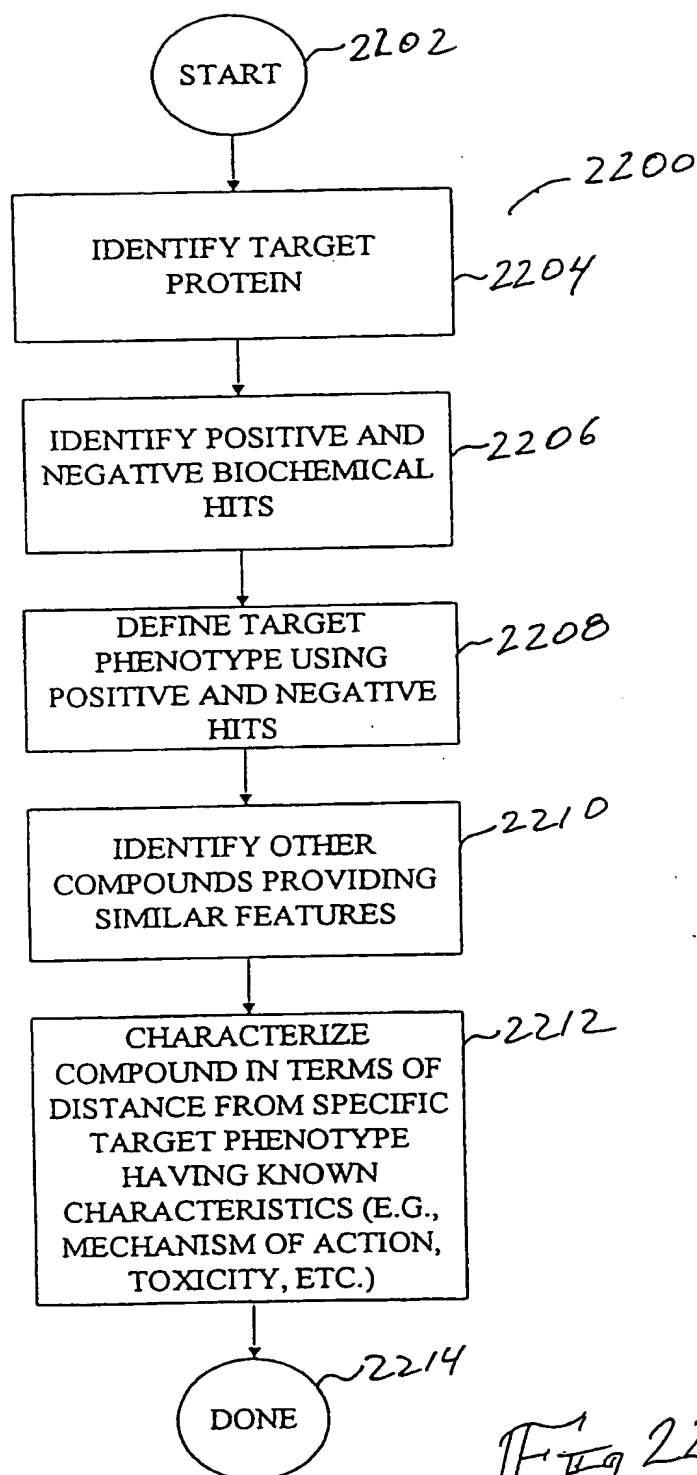


Fig 22